GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 11:26:35 ; Search time 546 Seconds
(without alignments)
2887.610 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 1577
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163412_14982/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10023597_@CGN_1_1_480_@runat_29092004_163412_14982
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1577	100.0	963	15	US-10-023-597-23	Sequence 23, Appl
2	1273.5	80.8	933	9	US-09-886-055-50	Sequence 50, Appl
3	1273.5	80.8	933	10	US-09-804-291-50	Sequence 50, Appl
4	1273.5	80.8	933	13	US-10-343-650A-395	Sequence 395, App
5	1273.5	80.8	933	15	US-10-220-382-35	Sequence 35, Appl
6	1273.5	80.8	1333	15	US-10-017-161-325	Sequence 325, App
7	1273.5	80.8	1333	16	US-10-292-798-291	Sequence 291, App
8	1268	80.4	976	15	US-10-024-399-1	Sequence 1, Appli
9	1266.5	80.3	953	16	US-10-005-041A-13	Sequence 13, Appl
10	1221	77.4	931	15	US-10-017-161-1301	Sequence 1301, Ap
11	1221	77.4	931	16	US-10-292-798-1069	Sequence 1069, Ap
12	1180	74.8	936	15	US-10-220-382-36	Sequence 36, Appl
13	1180	74.8	942	13	US-10-343-650A-397	Sequence 397, App
14	1180	74.8	951	16	US-10-387-629-47	Sequence 47, Appl
15	1180	74.8	1336	15	US-10-017-161-749	Sequence 749, App
16	1180	74.8	1336	16	US-10-292-798-653	Sequence 653, App
17	1058	67.1	952	15	US-10-023-597-59	Sequence 59, Appl
18	1048	66.5	961	15	US-10-023-601-89	Sequence 89, Appl
19	1024.5	65.0	934	15	US-10-023-597-125	Sequence 125, App
20	1023	64.9	981	15	US-10-023-597-81	Sequence 81, Appl
21	1013	64.2	943	15	US-10-023-597-57	Sequence 57, Appl
22	1006	63.8	950	15	US-10-023-597-87	Sequence 87, Appl
23	1006	63.8	953	15	US-10-023-597-63	Sequence 63, Appl
24	999.5	63.4	942	9	US-09-886-055-20	Sequence 20, Appl
25	999.5	63.4	942	10	US-09-804-291-20	Sequence 20, Appl
26	999.5	63.4	1342	15	US-10-017-161-321	Sequence 321, App
27	999.5	63.4	1342	16	US-10-292-798-287	Sequence 287, App
28	998.5	63.3	1342	15	US-10-017-161-315	Sequence 315, App
29	998.5	63.3	1342	16	US-10-292-798-281	Sequence 281, App
30	995.5	63.1	886	15	US-10-017-161-1299	Sequence 1299, Ap
31	995.5	63.1	886	16	US-10-292-798-1067	Sequence 1067, Ap
32	994.5	63.1	915	16	US-10-387-629-71	Sequence 71, Appl
33	992	62.9	981	15	US-10-023-597-91	Sequence 91, Appl
34	989	62.7	943	15	US-10-023-597-61	Sequence 61, Appl
35	978.5	62.0	942	16	US-10-387-629-73	Sequence 73, Appl
36	978.5	62.0	958	15	US-10-025-806-15	Sequence 15, Appl
37	978.5	62.0	974	15	US-10-024-399-35	Sequence 35, Appl
38	978.5	62.0	1342	15	US-10-017-161-313	Sequence 313, App
39	978.5	62.0	1342	16	US-10-292-798-279	Sequence 279, App
40	977	62.0	978	15	US-10-023-601-91	Sequence 91, Appl
41	977	62.0	1015	15	US-10-023-601-97	Sequence 97, Appl
42	976	61.9	942	15	US-10-023-601-99	Sequence 99, Appl
43	975	61.8	931	15	US-10-023-597-47	Sequence 47, Appl
44	972	61.6	943	15	US-10-023-597-89	Sequence 89, Appl
45	966.5	61.3	942	9	US-09-886-055-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-023-597-23
; Sequence 23, Application US/10023597
; Publication No. US20030109692A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Ballinger, Robert A.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20030109692A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-224AD
; CURRENT APPLICATION NUMBER: US/10/023,597
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,635
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/259,743
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/299,327
; PRIOR FILING DATE: 2001-06-19

```
;  PRIOR APPLICATION NUMBER: 60/261,498
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: 60/263,689
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: 60/267,464
;  PRIOR FILING DATE: 2001-02-08
;  PRIOR APPLICATION NUMBER: 60/271,021
;  PRIOR FILING DATE: 2001-02-22
;  PRIOR APPLICATION NUMBER: 60/275,946
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/278,150
;  PRIOR FILING DATE: 2001-03-23
;  PRIOR APPLICATION NUMBER: 60/285,718
;  PRIOR FILING DATE: 2001-04-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 963
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-023-597-23

Alignment Scores:
Pred. No.:                2.56e-146      Length:          963
Score:                    1577.00        Matches:         311
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       15             Gaps:            0

US-10-023-597-24 (1-311) x US-10-023-597-23 (1-963)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 61

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 121

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 181

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 241

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 301

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAG 361

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 421

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 481

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 541

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 601

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 661

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 721

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 781

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 841

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 901

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        902 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 934

RESULT 2
US-09-886-055-50
; Sequence 50, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
;  APPLICANT: STRYER, LUBERT
;  APPLICANT:  ZOZULYA, SERGEY
;  TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;  TITLE OF INVENTION:  BIOSENSORS OF CHEMICAL SENSANTS
;  FILE REFERENCE: 078003-0277150
;  CURRENT APPLICATION NUMBER: US/09/886,055
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 60/213,812
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 522
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-886-055-50

Alignment Scores:
Pred. No.:                2.59e-116      Length:          933
Score:                    1273.50        Matches:         250
Percent Similarity:       87.46%         Conservative:    22
Best Local Similarity:    80.39%         Mismatches:      38
Query Match:              80.75%         Indels:          1
DB:                       9              Gaps:            1

US-10-023-597-24 (1-311) x US-09-886-055-50 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||  ||||||||||||||||||||||||  ||||||  ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||::::::|||:::||||||||||||||||||||||||  ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||  ||||||||||||||||||  ||||||  |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||  ||||||||||||||:::||||||  ||||||||||||||  ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
```

```
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930
```

RESULT 3
US-09-804-291-50
; Sequence 50, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
; APPLICANT: ZOZULA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0278005
; CURRENT APPLICATION NUMBER: US/09/804,291
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291-50

Alignment Scores:

Pred. No.:	2.59e-116	Length:	933
Score:	1273.50	Matches:	250
Percent Similarity:	87.46%	Conservative:	22
Best Local Similarity:	80.39%	Mismatches:	38
Query Match:	80.75%	Indels:	1
DB:	10	Gaps:	1

US-10-023-597-24 (1-311) x US-09-804-291-50 (1-933)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
```

```
                ||||||  ||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::  ||||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||      |||      ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 4
US-10-343-650A-395
; Sequence 395, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
;  APPLICANT: HAGA, TATSUYA
;  TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
;  FILE REFERENCE: 31671-186347
;  CURRENT APPLICATION NUMBER: US/10/343,650A
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: JP 2000/237818
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: JP 2001/34434
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 694
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 395
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(933)
US-10-343-650A-395

Alignment Scores:
Pred. No.:              2.59e-116       Length:         933
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     13              Gaps:           1

US-10-023-597-24 (1-311) x US-10-343-650A-395 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||  |||||||||||||||||||||||  ||||||  ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::::|||:::|||||||||||||||||||||||||  ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||  ||||||||||||||||  ||||||  |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||  ||||||||||||||:::||||||  ||||||||||||||  ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||  ||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||  ||||||||||||||||||||||||:::|||||||||  ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||  ||||||||||||  ||||||:::||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::  ||||||||||||||||||||||||||||||||:::||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||  ||||||:::|||  ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||  |||  :::|||:::|||||||||  ||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||  ||||||  ||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||  ||||||||||||||||||||||||||||||  ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||  ||||||||||||||||:::||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::  ||||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||      |||      ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

*  RESULT 5
   US-10-220-382-35
   ; Sequence 35, Application US/10220382
   ; Publication No. US20030119111A1
   ; GENERAL INFORMATION:
   ;  APPLICANT: INCYTE GENOMICS, INC.
   ;  APPLICANT:  LAL, Preeti
   ;  APPLICANT:  TANG, Y. Tom
   ;  APPLICANT:  PATTERSON, Chandra
   ;  APPLICANT:  YAO, Monique G.
   ;  APPLICANT:  SHIH, Leo L.
   ;  APPLICANT:  TRIBOULEY, Catherine
   ;  APPLICANT:  LU, Dyung Aina M.
   ;  APPLICANT:  YUE, Henry
   ;  APPLICANT:  KHAN, Farrah A.
   ;  APPLICANT:  POLICKY, Jennifer L.
   ;  APPLICANT:  AU-YOUNG, Janice
   ;  APPLICANT:  YANG, Junming
   ;  APPLICANT:  HARLAND, Lee
   ;  APPLICANT:  WALSH, Roderick T.
   ;  APPLICANT:  LO, Terence P.
   ;  APPLICANT:  BOROWSKY, Mark L.
   ;  TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
   ;  FILE REFERENCE: PI-0044 PCT
   ;  CURRENT APPLICATION NUMBER: US/10/220,382
   ;  CURRENT FILING DATE:  2001-03-01
   ;  PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730
   ;  PRIOR FILING DATE: 2000-03-03; 2000-03-10; 2000-03-17; 2000-03-20
   ;  NUMBER OF SEQ ID NOS: 42
   ;  SOFTWARE: PERL Program
   ; SEQ ID NO 35
   ;   LENGTH: 933
   ;   TYPE: DNA
   ;   ORGANISM: Homo sapiens
   ;   FEATURE:
   ;   NAME/KEY: misc_feature
```

```
;   OTHER INFORMATION: Incyte ID No. US20030119111A1 7472439CB1
US-10-220-382-35

Alignment Scores:
Pred. No.:                2.59e-116      Length:          933
Score:                    1273.50        Matches:         250
Percent Similarity:       87.46%         Conservative:    22
Best Local Similarity:    80.39%         Mismatches:      38
Query Match:              80.75%         Indels:          1
DB:                       15             Gaps:            1

US-10-023-597-24 (1-311) x US-10-220-382-35 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||  |||||||||||||||||||||||||   ||||||  ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::||:::|||||||||||||||||||||||||  ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||  ||||||||||||||||||  ||||||  |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||  ||||||||||||||:::||||||  ||||||||||||||  ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 ||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||  ||||||||||||||||||||||||:::|||||||||  ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||  ||||||||||  ||||||:::||||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::  ||||||||||||||||||||||||||||||:::|||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||  ||||||:::|||  ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||  |||  :::|||:::||||||||||  ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||  ||||||  ||||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||  ||||||||||||||||||||||||||||  ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||  ||||||||||||||:::||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::  ||||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||  |||  ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 6
US-10-017-161-325
; Sequence 325, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 084335/0152
;  CURRENT APPLICATION NUMBER: US/10/017,161
;  CURRENT FILING DATE:  2002-12-18
;  PRIOR APPLICATION NUMBER: JP 2001/246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2430
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 325
;   LENGTH: 1333
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)..(1333)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(1133)
US-10-017-161-325

Alignment Scores:
Pred. No.:                4.32e-116      Length:          1333
Score:                    1273.50        Matches:         250
Percent Similarity:       87.46%         Conservative:    22
Best Local Similarity:    80.39%         Mismatches:      38
Query Match:              80.75%         Indels:          1
DB:                       15             Gaps:            1

US-10-023-597-24 (1-311) x US-10-017-161-325 (1-1333)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||  |||||||||||||||||||||||||   ||||||  ||||||
Db        201 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 257

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::||:::|||||||||||||||||||||||||  ||||||||||||
Db        258 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 317

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||  ||||||||||||||||||  ||||||  |||||||||
Db        318 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 377

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||  ||||||||||||||:::||||||  ||||||||||||||  ||||||
Db        378 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 437

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        438 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 497

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 ||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        498 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 557

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||  ||||||||||||||||||||||||:::|||||||||  ||||||
Db        558 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 617

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||  ||||||||||  ||||||:::||||||||||
```

```
Db        618 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 677

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        678 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 737

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        738 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 797

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        798 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 857

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        858 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 917

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        918 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 977

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        978 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 1037

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db       1038 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 1097

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db       1098 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 1130

RESULT 7
US-10-292-798-291
; Sequence 291, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 291
;   LENGTH: 1333
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   LOCATION: source
;   FEATURE:
;   LOCATION: (1)..(1333)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(1133)
US-10-292-798-291

Alignment Scores:
Pred. No.:              4.32e-116       Length:         1333
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     16              Gaps:           1
```

```
US-10-023-597-24 (1-311) x US-10-292-798-291 (1-1333)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db        201 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 257

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db        258 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 317

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        318 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 377

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db        378 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 437

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        438 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 497

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        498 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 557

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        558 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 617

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        618 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 677

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        678 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 737

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        738 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 797

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        798 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 857

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        858 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 917

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        918 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 977

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        978 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 1037

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db       1038 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 1097

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db       1098 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 1130

RESULT 8
US-10-024-399-1
; Sequence 1, Application US/10024399
; Publication No. US20030100491A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
```

```
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Colman, Steven D.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Ballinger, Robert A.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Li, Li
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Casman, Stacie J.
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-224AE
;  CURRENT APPLICATION NUMBER: US/10/024,399
;  CURRENT FILING DATE:  2001-12-18
;  PRIOR APPLICATION NUMBER: 60/256,635
;  PRIOR FILING DATE: 2000-12-18
;  PRIOR APPLICATION NUMBER: 60/259,743
;  PRIOR FILING DATE: 2001-01-04
;  PRIOR APPLICATION NUMBER: 60/299,327
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: 60/261,498
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: 60/263,689
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: 60/267,464
;  PRIOR FILING DATE: 2001-02-08
;  PRIOR APPLICATION NUMBER: 60/271,021
;  PRIOR FILING DATE: 2001-02-22
;  PRIOR APPLICATION NUMBER: 60/275,946
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/278,150
;  PRIOR FILING DATE: 2001-03-23
;  PRIOR APPLICATION NUMBER: 60/285,718
;  PRIOR FILING DATE: 2001-04-23
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 976
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-024-399-1

Alignment Scores:
Pred. No.:                9.69e-116      Length:          976
Score:                    1268.00        Matches:         247
Percent Similarity:       86.50%         Conservative:    22
Best Local Similarity:    79.42%         Mismatches:      42
Query Match:              80.41%         Indels:          0
DB:                       15             Gaps:            0

US-10-023-597-24 (1-311) x US-10-024-399-1 (1-976)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||           :::|||||||||||||||||||||||||   ||||||   ||||||
Db          8 ATGGGCCAGCCCAAAAACTCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 67

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::|||:::|||||||||||||||||||||||   ||||||||||||
Db         68 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 127

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        128 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 187

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        188 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 247

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        248 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 307

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||

Db        308 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 367

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||||||:::|||||||||   ||||||
Db        368 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 427

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||   ||||||||||||   ||||||:::|||||||||
Db        428 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 487

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||:::|||||||||
Db        488 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 547

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        548 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 607

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||||
Db        608 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 667

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||      ||||||   ||||||||||||||||||||||:::
Db        668 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 727

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||||   ||||||
Db        728 TGCAGTTCCCACATAATTGTAGTTTCTCTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 787

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::|||||||||||||||||||||||||||||||
Db        788 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 847

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::|||||:::   ||||||||||||||||||||||||||||||||||||||:::|||
Db        848 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 907

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||   |||   ||||||
Db        908 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 940

RESULT 9
US-10-005-041A-13
; Sequence 13, Application US/10005041A
; Publication No. US20030232331A1
; GENERAL INFORMATION:
;  APPLICANT: Casman, Stacie J
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Gilbert, Jennifer A
;  APPLICANT:  Mayotte, Jane E
;  APPLICANT:  Baumgartner, Jason C
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Vernet, Corine AM
;  APPLICANT:  Dickinson, Kevin S
;  APPLICANT:  Ballinger, Robert A
;  APPLICANT:  Wolenc, Adam R
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Gerlach, Valerie
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-215
;  CURRENT APPLICATION NUMBER: US/10/005,041A
;  CURRENT FILING DATE:  2001-12-04
;  PRIOR APPLICATION NUMBER: 60/251,459
;  PRIOR FILING DATE: 2000-12-05
```

```
;  PRIOR APPLICATION NUMBER: 60/259,007
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 205
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 953
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-005-041A-13

Alignment Scores:
Pred. No.:              1.32e-115       Length:         953
Score:                  1266.50         Matches:        249
Percent Similarity:     87.14%          Conservative:   22
Best Local Similarity:  80.06%          Mismatches:     39
Query Match:            80.31%          Indels:         1
DB:                     16              Gaps:           1

US-10-023-597-24 (1-311) x US-10-005-041A-13 (1-953)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||   ||||||   ||||||
Db         16 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 72

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::::|||:::|||||||||||||||||||||||   ||||||||||||
Db         73 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 132

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||   ||||||   |||||||||
Db        133 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 192

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        193 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 252

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        253 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 312

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||   ||||||||||||||||||||||||||
Db        313 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 372

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||:::||||||||   ||||||
Db        373 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 432

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||   ||||||||||||   ||||||:::||||||||||
Db        433 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 492

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||:::||||||||||
Db        493 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 552

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||   |||||:::|||   |||:::|||
Db        553 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCCGGTGGTCTTT 612

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||||
Db        613 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 672

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||      ||||||   ||||||||||||||||||||||:::
Db        673 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 732

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||:::||||||   ||||||||||||||||||||||||||||||   ||||||
Db        733 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 792

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
```

```
              ||||||   ||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        793 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 852

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||||||:::|||
Db        853 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 912

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||      |||      ||||||
Db        913 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 945

RESULT 10
US-10-017-161-1301
; Sequence 1301, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 084335/0152
;  CURRENT APPLICATION NUMBER: US/10/017,161
;  CURRENT FILING DATE:  2002-12-18
;  PRIOR APPLICATION NUMBER: JP 2001/246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2430
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1301
;   LENGTH: 931
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)..(931)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(731)
US-10-017-161-1301

Alignment Scores:
Pred. No.:              4.04e-111       Length:         931
Score:                  1221.00         Matches:        241
Percent Similarity:     99.59%          Conservative:   0
Best Local Similarity:  99.59%          Mismatches:     1
Query Match:            77.43%          Indels:         0
DB:                     15              Gaps:           0

US-10-023-597-24 (1-311) x US-10-017-161-1301 (1-931)

Qy         70 AspPheSerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLys 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAG 62

Qy         90 AsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPhe 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCTTT 122

Qy        110 SerGluSerPheIleLeuSerAlaMetValGluAspArgTyrValGlyIleCysAsnPro 129
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        123 TCTGAGTCCTTCATCCTGTCGGCGATGGTGTAGGACCGCTACGTGGGCATCTGTAACCCA 182

Qy        130 LeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyr 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTAC 242

Qy        150 GlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCys 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTGT 302

Qy        170 AlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCys 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGC 362
```

```
Qy        190 AsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyVal 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGGTG 422

Qy        210 ProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSer 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTAGT 482

Qy        230 SerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSer 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTTCT 542

Qy        250 LeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAsp 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGGAC 602

Qy        270 GlnGlyLysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIle 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAATC 662

Qy        290 TyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSer 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAAGC 722

Qy        310 PheSer 311
              ||||||
Db        723 TTTTCT 728

RESULT 11
US-10-292-798-1069
; Sequence 1069, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1069
;   LENGTH: 931
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   LOCATION: source
;   FEATURE:
;   LOCATION: (1)..(931)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(731)
US-10-292-798-1069

Alignment Scores:
Pred. No.:              4.04e-111       Length:         931
Score:                  1221.00         Matches:        241
Percent Similarity:     99.59%          Conservative:   0
Best Local Similarity:  99.59%          Mismatches:     1
Query Match:            77.43%          Indels:         0
DB:                     16              Gaps:           0

US-10-023-597-24 (1-311) x US-10-292-798-1069 (1-931)

Qy         70 AspPheSerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLys 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAG 62

Qy         90 AsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPhe 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCTTT 122

Qy        110 SerGluSerPheIleLeuSerAlaMetValGluAspArgTyrValGlyIleCysAsnPro 129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TCTGAGTCCTTCATCCTGTCGGCGATGGTGTAGGACCGCTACGTGGGCATCTGTAACCCA 182

Qy        130 LeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyr 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTAC 242

Qy        150 GlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCys 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTGT 302

Qy        170 AlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCys 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGC 362

Qy        190 AsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyVal 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGGTG 422

Qy        210 ProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSer 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTAGT 482

Qy        230 SerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSer 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTTCT 542

Qy        250 LeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAsp 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGGAC 602

Qy        270 GlnGlyLysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIle 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAATC 662

Qy        290 TyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSer 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAAGC 722

Qy        310 PheSer 311
              ||||||
Db        723 TTTTCT 728

RESULT 12
US-10-220-382-36
; Sequence 36, Application US/10220382
; Publication No. US20030119111A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  YAO, Monique G.
;  APPLICANT:  SHIH, Leo L.
;  APPLICANT:  TRIBOULEY, Catherine
;  APPLICANT:  LU, Dyung Aina M.
;  APPLICANT:  YUE, Henry
;  APPLICANT:  KHAN, Farrah A.
;  APPLICANT:  POLICKY, Jennifer L.
;  APPLICANT:  AU-YOUNG, Janice
;  APPLICANT:  YANG, Junming
;  APPLICANT:  HARLAND, Lee
;  APPLICANT:  WALSH, Roderick T.
;  APPLICANT:  LO, Terence P.
;  APPLICANT:  BOROWSKY, Mark L.
;  TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
```

```
;  FILE REFERENCE: PI-0044 PCT
;  CURRENT APPLICATION NUMBER: US/10/220,382
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730
;  PRIOR FILING DATE: 2000-03-03; 2000-03-10; 2000-03-17; 2000-03-20
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PERL Program
; SEQ ID NO 36
;   LENGTH: 936
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030119111A1 7472440CB1
US-10-220-382-36

Alignment Scores:
Pred. No.:              4.64e-107       Length:         936
Score:                  1180.00         Matches:        226
Percent Similarity:     83.92%          Conservative:   35
Best Local Similarity:  72.67%          Mismatches:     50
Query Match:            74.83%          Indels:         0
DB:                     15              Gaps:           0

US-10-023-597-24 (1-311) x US-10-220-382-36 (1-936)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||    ||||||:::|||||||||||||||||    |||
Db          1 ATGGCTGCTGAGAATTCCTCCTTCGTGACACAGTTTATCCTCGCAGGCTTAACTGACCAA 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||:::|||:::|||:::||||||||||||||||||||||||   ||||||||||||
Db         61 CCGGGAGTCCAGATCCCCCTCTTCTTCCTGTTTCTAGGCTTCTACGTGGTCACTGTGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||    ||||||    |||||||||   ||||||   |||||||||
Db        121 GGGAACCTGGGCTTGATAACCCTGATAAGGCTCAACTCTCACTTGCACACCCCTATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   :::||||||||||   :::||||||   :::|||:::   |||   ||||||
Db        181 TTCTTCCTCTATAACTTGTCCTTCATAGATTTCTGCTATTCCAGTGTTATCACTCCCAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||    :::||||||    ||||||:::    ||||||||||:::|||
Db        241 ATGCTGATGAGCTTTGTCTTAAAGAAGAACAGCATCTCCTACGCAGGGTGTATGACTCAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              |||||||||   |||||||||   ||||||||||||||||||||||||||||
Db        301 CTCTTCTTCTTTCTTTTCTTTGTTGTCTCTGAGTCCTTCATCCTGTCAGCAATGGCGTAT 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||    ||||||||||||||||||||||    :::||||||||||||||||||
Db        361 GACCGCTATGTGGCCATCTGTAACCCACTGTTGTACATGGTCACCATGTCTCCCCAGGTG 420

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||   ||||||||||||||||||||||||||||||       ||||||:::|||||||||
Db        421 TGTTTTCTCCTTTTGTTGGGTGTCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ::::::   :::||||||||||||:::||||||||||||||||||||||||||||||||
Db        481 GCGTGCATGATGGGTGTGACCTTCTGTGCCAATAACCTTGTCAACCACTACATGTGTGAC 540

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||    :::|||      :::|||:::|||   ||||||:::|||
Db        541 ATCCTTCCCCTTCTTGAGTGTGCTTGCACCAGCACCTATGTGAATGAGCTTGTAGTGTTT 600

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              :::||||||   :::   ||||||||||||   |||   :::||||||||||||
Db        601 GTTGTTGTGGGCATTGATATTGGTGTGCCCACAGTCACCATCTTCATTTCCTATGCTCTC 660

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||      :::   |||   ||||||||||||||||||||||:::
Db        661 ATTCTCTCCAGCATCTTCCACATTGATTCCACGGAGGGCAGGTCCAAAGCCTTCAGCACC 720

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||:::||||||||||||||||||||||||||||||||||||    ||||||
Db        721 TGCAGCTCCCACATAATTGCAGTTTCTCTGTTCTTTGGGTCAGGAGCATTCATGTACCTC 780

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   |||:::|||   ::::::||||||||||||||||||||||||||||||||
Db        781 AAACCCTTTTCTCTTTTAGCTATGAACCAGGGCAAGGTGTCTTCCCTATTCTATACCACT 840

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||   ||||||||||||||||||||||||||||||||||||:::|||
Db        841 GTGGTGCCCATGCTCAACCCATTAATTTATAGCCTGAGGAATAAGGACGTCAAAGTTGCT 900

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              ||||||:::      ::::::   :::|||||||
Db        901 CTAAAGAAAATCTTGAACAAAAATGCATTCTCC 933

RESULT 13
US-10-343-650A-397
; Sequence 397, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
;  APPLICANT: HAGA, TATSUYA
;  TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
;  FILE REFERENCE: 31671-186347
;  CURRENT APPLICATION NUMBER: US/10/343,650A
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: JP 2000/237818
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: JP 2001/34434
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 694
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 397
;   LENGTH: 942
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(942)
US-10-343-650A-397

Alignment Scores:
Pred. No.:              4.68e-107       Length:         942
Score:                  1180.00         Matches:        226
Percent Similarity:     83.92%          Conservative:   35
Best Local Similarity:  72.67%          Mismatches:     50
Query Match:            74.83%          Indels:         0
DB:                     13              Gaps:           0

US-10-023-597-24 (1-311) x US-10-343-650A-397 (1-942)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||    ||||||:::|||||||||||||||||||    |||
Db          7 ATGGCTGCTGAGAATTCCTCCTTCGTGACACAGTTTATCCTCGCAGGCTTAACTGACCAA 66

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||:::|||:::|||:::||||||||||||||||||||||   ||||||||||||
Db         67 CCGGGAGTCCAGATCCCCCTCTTCTTCCTGTTTCTAGGCTTCTACGTGGTCACTGTGGTG 126

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||    ||||||    |||||||||   ||||||   |||||||||
Db        127 GGGAACCTGGGCTTGATAACCCTGATAAGGCTCAACTCTCACTTGCACACCCCTATGTAC 186

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   :::||||||||||   :::||||||   :::|||:::   |||   ||||||
Db        187 TTCTTCCTCTATAACTTGTCCTTCATAGATTTCTGCTATTCCAGTGTTATCACTCCCAAA 246

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||    :::||||||    ||||||:::    ||||||||||:::|||
Db        247 ATGCTGATGAGCTTTGTCTTAAAGAAGAACAGCATCTCCTACGCAGGGTGTATGACTCAG 306

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
```

```
Db        307 CTCTTCTTCTTTCTTTTCTTTGTTGTCTCTGAGTCCTTCATCCTGTCAGCAATGGCGTAT 366

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        367 GACCGCTATGTGGCCATCTGTAACCCACTGTTGTACATGGTCACCATGTCTCCCCAGGTG 426

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        427 TGTTTTCTCCTTTTGTTGGGTGTCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 486

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        487 GCGTGCATGATGGGTGTGACCTTCTGTGCCAATAACCTTGTCAACCACTACATGTGTGAC 546

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        547 ATCCTTCCCCTTCTTGAGTGTGCTTGCACCAGCACCTATGTGAATGAGCTTGTAGTGTTT 606

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        607 GTTGTTGTGGGCATTGATATTGGTGTGCCCACAGTCACCATCTTCATTTCCTATGCTCTC 666

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        667 ATTCTCTCCAGCATCTTCCACATTGATTCCACGGAGGGCAGGTCCAAAGCCTTCAGCACC 726

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        727 TGCAGCTCCCACATAATTGCAGTTTCTCTGTTCTTTGGGTCAGGAGCATTCATGTACCTC 786

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        787 AAACCCTTTTCTCTTTTAGCTATGAACCAGGGCAAGGTGTCTTCCCTATTCTATACCACT 846

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        847 GTGGTGCCCATGCTCAACCCATTAATTTATAGCCTGAGGAATAAGGACGTCAAAGTTGCT 906

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db        907 CTAAAGAAAATCTTGAACAAAAATGCATTCTCC 939

RESULT 14
US-10-387-629-47
; Sequence 47, Application US/10387629
; Publication No. US20030221205A1
; GENERAL INFORMATION:
;  APPLICANT: ChemCom S.A.
;  APPLICANT:  Veithen, Alex
;  TITLE OF INVENTION: Olfactory and Pheromones G-Protein coupled Receptors
;  FILE REFERENCE: 9409/2192
;  CURRENT APPLICATION NUMBER: US/10/387,629
;  CURRENT FILING DATE:  2003-03-13
;  NUMBER OF SEQ ID NOS: 254
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
;   LENGTH: 951
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(948)
;   OTHER INFORMATION:
US-10-387-629-47

Alignment Scores:
Pred. No.:              4.74e-107       Length:         951
Score:                  1180.00         Matches:        226
Percent Similarity:     83.92%          Conservative:   35
Best Local Similarity:  72.67%          Mismatches:     50
Query Match:            74.83%          Indels:         0
DB:                     16              Gaps:           0

US-10-023-597-24 (1-311) x US-10-387-629-47 (1-951)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db         16 ATGGCTGCTGAGAATTCCTCCTTCGTGACACAGTTTATCCTCGCAGGCTTAACTGACCAA 75

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db         76 CCGGGAGTCCAGATCCCCCTCTTCTTCCTGTTTCTAGGCTTCTACGTGGTCACTGTGGTG 135

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        136 GGGAACCTGGGCTTGATAACCCTGATAAGGCTCAACTCTCACTTGCACACCCCTATGTAC 195

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db        196 TTCTTCCTCTATAACTTGTCCTTCATAGATTTCTGCTATTCCAGTGTTATCACTCCCAAA 255

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        256 ATGCTGATGAGCTTTGTCTTAAAGAAGAACAGCATCTCCTACGCAGGGTGTATGACTCAG 315

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        316 CTCTTCTTCTTTCTTTTCTTTGTTGTCTCTGAGTCCTTCATCCTGTCAGCAATGGCGTAT 375

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        376 GACCGCTATGTGGCCATCTGTAACCCACTGTTGTACATGGTCACCATGTCTCCCCAGGTG 435

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        436 TGTTTTCTCCTTTTGTTGGGTGTCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 495

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        496 GCGTGCATGATGGGTGTGACCTTCTGTGCCAATAACCTTGTCAACCACTACATGTGTGAC 555

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        556 ATCCTTCCCCTTCTTGAGTGTGCTTGCACCAGCACCTATGTGAATGAGCTTGTAGTGTTT 615

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        616 GTTGTTGTGGGCATTGATATTGGTGTGCCCACAGTCACCATCTTCATTTCCTATGCTCTC 675

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        676 ATTCTCTCCAGCATCTTCCACATTGATTCCACGGAGGGCAGGTCCAAAGCCTTCAGCACC 735

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        736 TGCAGCTCCCACATAATTGCAGTTTCTCTGTTCTTTGGGTCAGGAGCATTCATGTACCTC 795

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        796 AAACCCTTTTCTCTTTTAGCTATGAACCAGGGCAAGGTGTCTTCCCTATTCTATACCACT 855

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        856 GTGGTGCCCATGCTCAACCCATTAATTTATAGCCTGAGGAATAAGGACGTCAAAGTTGCT 915

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db        916 CTAAAGAAAATCTTGAACAAAAATGCATTCTCC 948

RESULT 15
US-10-017-161-749
; Sequence 749, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
```

```
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 084335/0152
;  CURRENT APPLICATION NUMBER: US/10/017,161
;  CURRENT FILING DATE:  2002-12-18
;  PRIOR APPLICATION NUMBER: JP 2001/246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2430
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 749
;   LENGTH: 1336
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)..(1336)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(1136)
US-10-017-161-749

Alignment Scores:
Pred. No.:              7.72e-107      Length:          1336
Score:                  1180.00        Matches:         226
Percent Similarity:     83.92%         Conservative:    35
Best Local Similarity:  72.67%         Mismatches:      50
Query Match:            74.83%         Indels:          0
DB:                     15             Gaps:            0

US-10-023-597-24 (1-311) x US-10-017-161-749 (1-1336)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||    ||||||:::|||||||||||||||||||      |||
Db        201 ATGGCTGCTGAGAATTCCTCCTTCGTGACACAGTTTATCCTCGCAGGCTTAACTGACCAA 260

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||:::|||:::|||:::|||||||||||||||||||||||||   ||||||||||||
Db        261 CCGGGAGTCCAGATCCCCCTCTTCTTCCTGTTTCTAGGCTTCTACGTGGTCACTGTGGTG 320

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||  ||||||||  ||||||  |||||||||
Db        321 GGGAACCTGGGCTTGATAACCCTGATAAGGCTCAACTCTCACTTGCACACCCCTATGTAC 380

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   :::||||||||||   :::|||||||   :::|||:::   |||   ||||||
Db        381 TTCTTCCTCTATAACTTGTCCTTCATAGATTTCTGCTATTCCAGTGTTATCACTCCCAAA 440

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||   :::|||||||   ||||||:::   |||||||||:::|||
Db        441 ATGCTGATGAGCTTTGTCTTAAAGAAGAACAGCATCTCCTACGCAGGGTGTATGACTCAG 500

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 |||||||||   |||||||||   ||||||||||||||||||||||||
Db        501 CTCTTCTTCTTTCTTTTCTTTGTTGTCTCTGAGTCCTTCATCCTGTCAGCAATGGCGTAT 560

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||   :::|||||||||||||||||||
Db        561 GACCGCTATGTGGCCATCTGTAACCCACTGTTGTACATGGTCACCATGTCTCCCCAGGTG 620

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||   ||||||||||||||||||||||||||||||||||||||      ||||||:::|||||||||
Db        621 TGTTTTCTCCTTTTGTTGGGTGTCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 680

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
                  ::::::   :::||||||||||||:::|||||||||||||||||||||||||||||
Db        681 GCGTGCATGATGGGTGTGACCTTCTGTGCCAATAACCTTGTCAACCACTACATGTGTGAC 740

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||   :::|||      :::|||:::|||   ||||||:::|||
Db        741 ATCCTTCCCCTTCTTGAGTGTGCTTGCACCAGCACCTATGTGAATGAGCTTGTAGTGTTT 800

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              :::||||||   :::   ||||||||||||   |||   :::||||||||||||
Db        801 GTTGTTGTGGGCATTGATATTGGTGTGCCCACAGTCACCATCTTCATTTCCTATGCTCTC 860

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||      :::   |||   ||||||||||||||||||||||:::
Db        861 ATTCTCTCCAGCATCTTCCACATTGATTCCACGGAGGGCAGGTCCAAAGCCTTCAGCACC 920

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||||||||||||||||||||||||||||||||||||   ||||||
Db        921 TGCAGCTCCCACATAATTGCAGTTTCTCTGTTCTTTGGGTCAGGAGCATTCATGTACCTC 980

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   |||:::|||   ::::::||||||||||||||||||||||||||||||||
Db        981 AAACCCTTTTCTCTTTTAGCTATGAACCAGGGCAAGGTGTCTTCCCTATTCTATACCACT 1040

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||:::|||
Db       1041 GTGGTGCCCATGCTCAACCCATTAATTTATAGCCTGAGGAATAAGGACGTCAAAGTTGCT 1100

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              ||||||:::      ::::::   :::||||||
Db       1101 CTAAAGAAAATCTTGAACAAAAATGCATTCTCC 1133


Search completed: September 30, 2004, 13:55:03
Job time : 553 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 07:23:04 ; Search time 561 Seconds
(without alignments)
8702.305 Million cell updates/sec

Title: US-10-023-597-23
Perfect score: 963
Sequence: 1 aatggctgccgagaactcct.....ttttagaaacagaaaagaga 963

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3340653 seqs, 2534783454 residues

Word size : 20

Total number of hits satisfying chosen parameters: 153

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	963	100.0	963	15	US-10-023-597-23	Sequence 23, Appl
2	706	73.3	931	15	US-10-017-161-1301	Sequence 1301, Ap
3	706	73.3	931	16	US-10-292-798-1069	Sequence 1069, Ap
4	121	12.6	886	15	US-10-017-161-1299	Sequence 1299, Ap
5	121	12.6	886	16	US-10-292-798-1067	Sequence 1067, Ap
6	57	5.9	487	9	US-09-747-155-366	Sequence 366, App
7	39	4.0	486	9	US-09-747-155-359	Sequence 359, App
8	39	4.0	487	9	US-09-747-155-299	Sequence 299, App
9	39	4.0	487	9	US-09-747-155-348	Sequence 348, App
10	39	4.0	489	9	US-09-747-155-279	Sequence 279, App
11	39	4.0	491	9	US-09-747-155-313	Sequence 313, App
12	39	4.0	933	9	US-09-886-055-50	Sequence 50, Appl
13	39	4.0	933	10	US-09-804-291-50	Sequence 50, Appl
14	39	4.0	933	13	US-10-343-650A-395	Sequence 395, App
15	39	4.0	933	15	US-10-220-382-35	Sequence 35, Appl
16	39	4.0	953	16	US-10-005-041A-13	Sequence 13, Appl
17	39	4.0	976	15	US-10-024-399-1	Sequence 1, Appli
18	39	4.0	1333	15	US-10-017-161-325	Sequence 325, App
19	39	4.0	1333	16	US-10-292-798-291	Sequence 291, App
c 20	34	3.5	34	16	US-10-085-198-230	Sequence 230, App
21	33	3.4	485	9	US-09-747-155-232	Sequence 232, App
22	33	3.4	486	9	US-09-747-155-233	Sequence 233, App
23	33	3.4	936	15	US-10-220-382-36	Sequence 36, Appl
24	33	3.4	942	13	US-10-343-650A-397	Sequence 397, App
25	33	3.4	951	16	US-10-387-629-47	Sequence 47, Appl
26	33	3.4	1336	15	US-10-017-161-749	Sequence 749, App
27	33	3.4	1336	16	US-10-292-798-653	Sequence 653, App
28	30	3.1	945	16	US-10-300-846-13	Sequence 13, Appl
29	28	2.9	945	15	US-10-023-597-95	Sequence 95, Appl
30	28	2.9	950	15	US-10-023-597-113	Sequence 113, App
31	28	2.9	962	15	US-10-023-597-115	Sequence 115, App
32	26	2.7	951	9	US-09-886-055-134	Sequence 134, App
33	26	2.7	951	10	US-09-804-291-134	Sequence 134, App
34	26	2.7	951	13	US-10-343-650A-283	Sequence 283, App
35	26	2.7	951	16	US-10-387-629-21	Sequence 21, Appl
36	26	2.7	971	13	US-10-433-581-21	Sequence 21, Appl
37	26	2.7	980	15	US-10-025-806-49	Sequence 49, Appl
38	26	2.7	980	15	US-10-025-806-51	Sequence 51, Appl
39	26	2.7	1013	15	US-10-023-601-85	Sequence 85, Appl
40	26	2.7	1351	15	US-10-017-161-221	Sequence 221, App
41	26	2.7	1351	15	US-10-017-161-331	Sequence 331, App
42	26	2.7	1351	16	US-10-292-798-191	Sequence 191, App
43	23	2.4	23	16	US-10-085-198-229	Sequence 229, App
44	23	2.4	507	10	US-09-777-789-45	Sequence 45, Appl
45	23	2.4	528	10	US-09-777-789-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-10-023-597-23
; Sequence 23, Application US/10023597
; Publication No. US20030109692A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Ballinger, Robert A.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20030109692A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-224AD
; CURRENT APPLICATION NUMBER: US/10/023,597
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,635
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/259,743
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/299,327
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 60/261,498
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/263,689
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/267,464
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/271,021
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/275,946
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/278,150
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/285,718
; PRIOR FILING DATE: 2001-04-03

```
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 963
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-023-597-23

  Query Match             100.0%;  Score 963;  DB 15;  Length 963;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  963;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCA 60

Qy         61 GCCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGT 120

Qy        121 GGGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTA 180

Qy        181 CTTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAA 240

Qy        241 AATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCA 300

Qy        301 GTTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGA 360

Qy        361 GGACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGT 420

Qy        421 GTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATAC 480

Qy        481 AGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGA 540

Qy        541 CATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTT 600

Qy        601 TATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTT 660

Qy        661 TATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAG 720

Qy        721 CTGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCT 780

Qy        781 CAAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCAC 840

Qy        841 TGTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGC 900

Qy        901 CCTGAAGAGAACCTTTTCCAGAATAAGCTTTTCTTGAAAAAAATTTTAGAAACAGAAAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTGAAGAGAACCTTTTCCAGAATAAGCTTTTCTTGAAAAAAATTTTAGAAACAGAAAAG 960

Qy        961 AGA 963
              |||
Db        961 AGA 963


RESULT 2
US-10-017-161-1301
; Sequence 1301, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 084335/0152
;  CURRENT APPLICATION NUMBER: US/10/017,161
;  CURRENT FILING DATE:  2002-12-18
;  PRIOR APPLICATION NUMBER: JP 2001/246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2430
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1301
;   LENGTH: 931
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)..(931)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(731)
US-10-017-161-1301

  Query Match             73.3%;  Score 706;  DB 15;  Length 931;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  756;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        207 TAGATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGA 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGA 60

Qy        267 AGAACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCT 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCT 120

Qy        327 TTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAGGACCGCTACGTGGGCATCTGTAACC 386
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTAGGACCGCTACGTGGGCATCTGTAACC 180

Qy        387 CACTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCT 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCT 240

Qy        447 ACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTT 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTT 300

Qy        507 GTGCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCT 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCT 360

Qy        567 GCAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGG 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGG 420

Qy        627 TGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTA 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTA 480
```

```
Qy        687 GTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTT 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTT 540

Qy        747 CTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGG 806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGG 600

Qy        807 ACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAA 866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAA 660

Qy        867 TCTACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAA 926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCTACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAA 720

Qy        927 GCTTTTCTTGAAAAAAATTTTAGAAACAGAAAAGAGA 963
              |||||||||||||||||||||||||||||||||||||
Db        721 GCTTTTCTTGAAAAAAATTTTAGAAACAGAAAAGAGA 757


RESULT 3
US-10-292-798-1069
; Sequence 1069, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1069
;   LENGTH: 931
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   LOCATION: source
;   FEATURE:
;   LOCATION: (1)..(931)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(731)
US-10-292-798-1069

  Query Match           73.3%;  Score 706;  DB 16;  Length 931;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  756;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        207 TAGATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGA 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGA 60

Qy        267 AGAACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCT 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCT 120

Qy        327 TTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAGGACCGCTACGTGGGCATCTGTAACC 386
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 TTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTAGGACCGCTACGTGGGCATCTGTAACC 180

Qy        387 CACTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCT 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCT 240

Qy        447 ACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTT 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTT 300

Qy        507 GTGCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCT 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCT 360

Qy        567 GCAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGG 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGG 420

Qy        627 TGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTA 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTA 480

Qy        687 GTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTT 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTT 540

Qy        747 CTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGG 806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGG 600

Qy        807 ACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAA 866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAA 660

Qy        867 TCTACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAA 926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCTACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAA 720

Qy        927 GCTTTTCTTGAAAAAAATTTTAGAAACAGAAAAGAGA 963
              |||||||||||||||||||||||||||||||||||||
Db        721 GCTTTTCTTGAAAAAAATTTTAGAAACAGAAAAGAGA 757


RESULT 4
US-10-017-161-1299
; Sequence 1299, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 084335/0152
;  CURRENT APPLICATION NUMBER: US/10/017,161
;  CURRENT FILING DATE:  2002-12-18
;  PRIOR APPLICATION NUMBER: JP 2001/246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2430
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1299
;   LENGTH: 886
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)..(886)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(686)
US-10-017-161-1299

  Query Match           12.6%;  Score 121;  DB 15;  Length 886;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-53;
  Matches  121;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         82 CTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAAT 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        281 CTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAAT 340

Qy        142 CCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTC 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 CCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTC 400

Qy        202 C 202
              |
Db        401 C 401


RESULT 5
US-10-292-798-1067
; Sequence 1067, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1067
;   LENGTH: 886
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   LOCATION: source
;   FEATURE:
;   LOCATION: (1)..(886)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(686)
US-10-292-798-1067

  Query Match             12.6%;  Score 121;  DB 16;  Length 886;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-53;
  Matches  121;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         82 CTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAAT 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 CTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAAT 340

Qy        142 CCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTC 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 CCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTC 400

Qy        202 C 202
              |
Db        401 C 401


RESULT 6
US-09-747-155-366
; Sequence 366, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
;  PRIOR FILING DATE: 1999-12-22
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 366
;   LENGTH: 487
;   TYPE: DNA
;   ORGANISM: Saimiri boliviensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(487)
;   OTHER INFORMATION: Taxon = 27679; gene = SBO222; Accession DDBJ/EMBL/GenBank = AF179
;   NAME/KEY: CDS
;   LOCATION: (2)..(487)
;   OTHER INFORMATION: Product = olfactory receptor
US-09-747-155-366

  Query Match             5.9%;  Score 57;  DB 9;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-19;
  Matches   57;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        514 CAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAA 570
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 CAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAA 201


RESULT 7
US-09-747-155-359
; Sequence 359, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
;  PRIOR FILING DATE: 1999-12-22
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 359
;   LENGTH: 486
;   TYPE: DNA
;   ORGANISM: Saimiri boliviensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(486)
;   OTHER INFORMATION: Taxon = 27679; gene = SBO218; Accession DDBJ/EMBL/GenBank = AF179
;   NAME/KEY: CDS
;   LOCATION: (2)..(484)
;   OTHER INFORMATION: Product = olfactory receptor
US-09-747-155-359

  Query Match             4.0%;  Score 39;  DB 9;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        642 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 680
              |||||||||||||||||||||||||||||||||||||||
Db        273 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 311


RESULT 8
US-09-747-155-299
; Sequence 299, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
```

```
;  PRIOR FILING DATE: 1999-12-22
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 299
;   LENGTH: 487
;   TYPE: DNA
;   ORGANISM: Callithrix jacchus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(487)
;   OTHER INFORMATION: Taxon = 9483; gene = CJA171; Accession DDBJ/EMBL/GenBank = AF1798
;   NAME/KEY: CDS
;   LOCATION: (2)..(487)
;   OTHER INFORMATION: Product = olfactory receptor
US-09-747-155-299

  Query Match             4.0%;  Score 39;  DB 9;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        642 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 680
              |||||||||||||||||||||||||||||||||||||||
Db        273 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 311


RESULT 9
US-09-747-155-348
; Sequence 348, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
;  PRIOR FILING DATE: 1999-12-22
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 348
;   LENGTH: 487
;   TYPE: DNA
;   ORGANISM: Saimiri sciureus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(487)
;   OTHER INFORMATION: Taxon = 9521; gene = SSC195; Accession DDBJ/EMBL/GenBank = AF1798
;   NAME/KEY: CDS
;   LOCATION: (2)..(487)
;   OTHER INFORMATION: Product = olfactory receptor
US-09-747-155-348

  Query Match             4.0%;  Score 39;  DB 9;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        642 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 680
              |||||||||||||||||||||||||||||||||||||||
Db        273 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 311


RESULT 10
US-09-747-155-279
; Sequence 279, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
;  PRIOR FILING DATE: 1999-12-22
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 279
;   LENGTH: 489
;   TYPE: DNA
;   ORGANISM: Macaca sylvanus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(489)
;   OTHER INFORMATION: Taxon = 9546; gene = MSY173; pseudogene; Accession DDBJ/EMBL/GenB
US-09-747-155-279

  Query Match             4.0%;  Score 39;  DB 9;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        642 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 680
              |||||||||||||||||||||||||||||||||||||||
Db        273 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 311


RESULT 11
US-09-747-155-313
; Sequence 313, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
;  PRIOR FILING DATE: 1999-12-22
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 313
;   LENGTH: 491
;   TYPE: DNA
;   ORGANISM: Pongo pygmaeus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(491)
;   OTHER INFORMATION: Taxon = 9600; gene = PPY110; pseudogene; Accession DDBJ/EMBL/GenB
US-09-747-155-313

  Query Match             4.0%;  Score 39;  DB 9;  Length 491;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        642 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 680
              |||||||||||||||||||||||||||||||||||||||
Db        273 TTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCC 311


RESULT 12
US-09-886-055-50
; Sequence 50, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
;  APPLICANT: STRYER, LUBERT
;  APPLICANT:  ZOZULYA, SERGEY
;  TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;  TITLE OF INVENTION:  BIOSENSORS OF CHEMICAL SENSANTS
;  FILE REFERENCE: 078003-0277150
;  CURRENT APPLICATION NUMBER: US/09/886,055
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 60/213,812
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 522
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 50
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-886-055-50

  Query Match             4.0%;  Score 39;  DB 9;  Length 933;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        235 TCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACAT 273
              |||||||||||||||||||||||||||||||||||||||
Db        231 TCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACAT 269


RESULT 13
US-09-804-291-50
; Sequence 50, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
;  APPLICANT: ZOZULA, SERGEY
;  TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
;  FILE REFERENCE: P 0278005
;  CURRENT APPLICATION NUMBER: US/09/804,291
;  CURRENT FILING DATE:  2001-03-13
;  PRIOR APPLICATION NUMBER: 60/188,914
;  PRIOR FILING DATE: 2000-03-13
;  PRIOR APPLICATION NUMBER: 60/192,033
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/198,474
;  PRIOR FILING DATE: 2000-04-14
;  PRIOR APPLICATION NUMBER: 60/199,335
;  PRIOR FILING DATE: 2000-04-24
;  PRIOR APPLICATION NUMBER: 60/207,702
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/213,849
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: 60/226,534
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: 60/230,732
;  PRIOR FILING DATE: 2000-09-07
;  PRIOR APPLICATION NUMBER: 60/266,862
;  PRIOR FILING DATE: 2001-02-07
;  NUMBER OF SEQ ID NOS: 529
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-804-291-50

  Query Match             4.0%;  Score 39;  DB 10;  Length 933;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        235 TCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACAT 273
              |||||||||||||||||||||||||||||||||||||||
Db        231 TCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACAT 269


RESULT 14
US-10-343-650A-395
; Sequence 395, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
;  APPLICANT: HAGA, TATSUYA
;  TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
;  FILE REFERENCE: 31671-186347
;  CURRENT APPLICATION NUMBER: US/10/343,650A
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: JP 2000/237818
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: JP 2001/34434
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 694
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 395
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(933)
US-10-343-650A-395

  Query Match             4.0%;  Score 39;  DB 13;  Length 933;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        235 TCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACAT 273
              |||||||||||||||||||||||||||||||||||||||
Db        231 TCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACAT 269


*  RESULT 15
US-10-220-382-35
; Sequence 35, Application US/10220382
; Publication No. US20030119111A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  YAO, Monique G.
;  APPLICANT:  SHIH, Leo L.
;  APPLICANT:  TRIBOULEY, Catherine
;  APPLICANT:  LU, Dyung Aina M.
;  APPLICANT:  YUE, Henry
;  APPLICANT:  KHAN, Farrah A.
;  APPLICANT:  POLICKY, Jennifer L.
;  APPLICANT:  AU-YOUNG, Janice
;  APPLICANT:  YANG, Junming
;  APPLICANT:  HARLAND, Lee
;  APPLICANT:  WALSH, Roderick T.
;  APPLICANT:  LO, Terence P.
;  APPLICANT:  BOROWSKY, Mark L.
;  TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: PI-0044 PCT
;  CURRENT APPLICATION NUMBER: US/10/220,382
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730
;  PRIOR FILING DATE: 2000-03-03; 2000-03-10; 2000-03-17; 2000-03-20
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PERL Program
; SEQ ID NO 35
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030119111A1 7472439CB1
US-10-220-382-35

  Query Match             4.0%;  Score 39;  DB 15;  Length 933;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        235 TCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACAT 273
              |||||||||||||||||||||||||||||||||||||||
Db        231 TCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACAT 269


Search completed: September 30, 2004, 10:01:34
Job time : 562 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 15:18:29 ; Search time 546 Seconds
(without alignments)
2887.610 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 311
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3340653 seqs, 2534783454 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6672343

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163433_15514/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=olig.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10023597_@CGN_1_1_480_@runat_29092004_163433_15514
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	311	100.0	963	15	US-10-023-597-23	Sequence 23, Appl
	2	191	61.4	931	15	US-10-017-161-1301	Sequence 1301, Ap
	3	191	61.4	931	16	US-10-292-798-1069	Sequence 1069, Ap
	4	40	12.9	886	15	US-10-017-161-1299	Sequence 1299, Ap
	5	40	12.9	886	16	US-10-292-798-1067	Sequence 1067, Ap
	6	33	10.6	487	9	US-09-747-155-366	Sequence 366, App
	7	20	6.4	485	9	US-09-747-155-232	Sequence 232, App
	8	20	6.4	933	9	US-09-886-055-50	Sequence 50, Appl
	9	20	6.4	933	10	US-09-804-291-50	Sequence 50, Appl
	10	20	6.4	933	13	US-10-343-650A-395	Sequence 395, App
	11	20	6.4	933	15	US-10-220-382-35	Sequence 35, Appl
	12	20	6.4	953	16	US-10-005-041A-13	Sequence 13, Appl
	13	20	6.4	976	15	US-10-024-399-1	Sequence 1, Appli
	14	20	6.4	1333	15	US-10-017-161-325	Sequence 325, App
	15	20	6.4	1333	16	US-10-292-798-291	Sequence 291, App
	16	19	6.1	486	9	US-09-747-155-359	Sequence 359, App
	17	19	6.1	487	9	US-09-747-155-348	Sequence 348, App
	18	19	6.1	958	15	US-10-023-597-67	Sequence 67, Appl
	19	19	6.1	993	15	US-10-023-601-93	Sequence 93, Appl
	20	18	5.8	487	9	US-09-747-155-299	Sequence 299, App
	21	18	5.8	1015	15	US-10-023-601-97	Sequence 97, Appl
	22	17	5.5	486	9	US-09-747-155-233	Sequence 233, App
	23	17	5.5	489	9	US-09-747-155-279	Sequence 279, App
	24	17	5.5	871	15	US-10-017-161-1383	Sequence 1383, Ap
	25	17	5.5	871	16	US-10-292-798-1129	Sequence 1129, Ap
	26	17	5.5	937	15	US-10-025-806-13	Sequence 13, Appl
	27	16	5.1	934	15	US-10-023-597-125	Sequence 125, App
	28	16	5.1	953	15	US-10-023-597-63	Sequence 63, Appl
c	29	15	4.8	554	15	US-10-029-386-3264	Sequence 3264, Ap
	30	15	4.8	801	13	US-10-343-650A-431	Sequence 431, App
	31	15	4.8	933	16	US-10-387-629-251	Sequence 251, App
	32	15	4.8	936	15	US-10-220-382-36	Sequence 36, Appl
	33	15	4.8	942	13	US-10-343-650A-397	Sequence 397, App
	34	15	4.8	945	15	US-10-023-597-127	Sequence 127, App
	35	15	4.8	951	16	US-10-387-629-47	Sequence 47, Appl
	36	15	4.8	955	16	US-10-024-212-7	Sequence 7, Appli
	37	15	4.8	955	16	US-10-024-212-9	Sequence 9, Appli
	38	15	4.8	955	16	US-10-024-212-11	Sequence 11, Appl
	39	15	4.8	955	16	US-10-024-212-13	Sequence 13, Appl
	40	15	4.8	955	16	US-10-024-212-15	Sequence 15, Appl
	41	15	4.8	974	15	US-10-024-399-35	Sequence 35, Appl
	42	15	4.8	1333	15	US-10-017-161-545	Sequence 545, App
	43	15	4.8	1333	16	US-10-292-798-477	Sequence 477, App
	44	15	4.8	1334	13	US-10-182-822A-36	Sequence 36, Appl
	45	15	4.8	1334	15	US-10-219-834-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-10-023-597-23
; Sequence 23, Application US/10023597
; Publication No. US20030109692A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Ballinger, Robert A.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20030109692A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-224AD
; CURRENT APPLICATION NUMBER: US/10/023,597
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,635
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/259,743
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/299,327
; PRIOR FILING DATE: 2001-06-19

```
;  PRIOR APPLICATION NUMBER: 60/261,498
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: 60/263,689
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: 60/267,464
;  PRIOR FILING DATE: 2001-02-08
;  PRIOR APPLICATION NUMBER: 60/271,021
;  PRIOR FILING DATE: 2001-02-22
;  PRIOR APPLICATION NUMBER: 60/275,946
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/278,150
;  PRIOR FILING DATE: 2001-03-23
;  PRIOR APPLICATION NUMBER: 60/285,718
;  PRIOR FILING DATE: 2001-04-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 963
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-023-597-23

Alignment Scores:
Pred. No.:                5.24e-300      Length:          963
Score:                    311.00         Matches:         311
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       15             Gaps:            0

US-10-023-597-24 (1-311) x US-10-023-597-23 (1-963)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 61

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 121

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 181

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 241

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 301

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAG 361

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 421

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 481

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 541

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 601

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 661

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 721

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 781

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 841

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 901

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        902 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 934

RESULT 2
US-10-017-161-1301
; Sequence 1301, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 084335/0152
;  CURRENT APPLICATION NUMBER: US/10/017,161
;  CURRENT FILING DATE:  2002-12-18
;  PRIOR APPLICATION NUMBER: JP 2001/246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2430
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1301
;   LENGTH: 931
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)..(931)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(731)
US-10-017-161-1301

Alignment Scores:
Pred. No.:                2.15e-180      Length:          931
Score:                    191.00         Matches:         241
Percent Similarity:       99.18%         Conservative:    0
Best Local Similarity:    99.18%         Mismatches:      1
Query Match:              61.41%         Indels:          2
DB:                       15             Gaps:            0

US-10-023-597-24 (1-311) x US-10-017-161-1301 (1-931)

Qy         70 AspPheSerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLys 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAG 62

Qy         90 AsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPhe 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCTTT 122

Qy        110 SerGluSerPheIleLeuSerAlaMetValGlu-AspArgTyrValGlyIleCysAsnPr 129
              |||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||
```

```
Db        123 TCTGAGTCCTTCATCCTGTCGGCGATGGTGTA-GGACCGCTACGTGGGCATCTGTAACCC 181

Qy        129 oLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTy 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTA 241

Qy        149 rGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCy 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTG 301

Qy        169 sAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCy 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTG 361

Qy        189 sAsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyVa 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGGT 421

Qy        209 lProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSe 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GCCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTAG 481

Qy        229 rSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSe 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTTC 541

Qy        249 rLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAs 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGGA 601

Qy        269 pGlnGlyLysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIl 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CCAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAAT 661

Qy        289 eTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSe 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CTACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAAG 721

Qy        309 rPheSer 311
              |||||||
Db        722 CTTTTCT 728
```

```
RESULT 3
US-10-292-798-1069
; Sequence 1069, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1069
;   LENGTH: 931
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   LOCATION: source
;   FEATURE:
;   LOCATION: (1)..(931)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(731)
US-10-292-798-1069
```

```
Alignment Scores:
Pred. No.:              2.15e-180       Length:          931
Score:                  191.00          Matches:         241
Percent Similarity:     99.18%          Conservative:    0
Best Local Similarity:  99.18%          Mismatches:      1
Query Match:            61.41%          Indels:          2
DB:                     16              Gaps:            0

US-10-023-597-24 (1-311) x US-10-292-798-1069 (1-931)

Qy         70 AspPheSerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLys 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAG 62

Qy         90 AsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPhe 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCTTT 122

Qy        110 SerGluSerPheIleLeuSerAlaMetValGlu-AspArgTyrValGlyIleCysAsnPr 129
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        123 TCTGAGTCCTTCATCCTGTCGGCGATGGTGTA-GGACCGCTACGTGGGCATCTGTAACCC 181

Qy        129 oLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTy 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTA 241

Qy        149 rGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCy 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTG 301

Qy        169 sAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCy 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTG 361

Qy        189 sAsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyVa 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGGT 421

Qy        209 lProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSe 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GCCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTAG 481

Qy        229 rSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSe 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTTC 541

Qy        249 rLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAs 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGGA 601

Qy        269 pGlnGlyLysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIl 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CCAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAAT 661

Qy        289 eTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSe 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CTACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAAG 721

Qy        309 rPheSer 311
              |||||||
Db        722 CTTTTCT 728
```

```
RESULT 4
US-10-017-161-1299
; Sequence 1299, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
```

```
;  FILE REFERENCE: 084335/0152
;  CURRENT APPLICATION NUMBER: US/10/017,161
;  CURRENT FILING DATE:  2002-12-18
;  PRIOR APPLICATION NUMBER: JP 2001/246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2430
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1299
;   LENGTH: 886
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)..(886)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(686)
US-10-017-161-1299

Alignment Scores:
Pred. No.:              6.97e-30        Length:         886
Score:                  40.00           Matches:        40
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            12.86%          Indels:         0
DB:                     15              Gaps:           0

US-10-023-597-24 (1-311) x US-10-017-161-1299 (1-886)

Qy         28 PhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIle 47
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 TTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAATC 341

Qy         48 LeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSer 67
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTCC 401

RESULT 5
US-10-292-798-1067
; Sequence 1067, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1067
;   LENGTH: 886
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   LOCATION: source
;   FEATURE:
;   LOCATION: (1)..(886)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(686)
US-10-292-798-1067

Alignment Scores:
Pred. No.:              6.97e-30        Length:         886
Score:                  40.00           Matches:        40
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            12.86%          Indels:         0
DB:                     16              Gaps:           0

US-10-023-597-24 (1-311) x US-10-292-798-1067 (1-886)

Qy         28 PhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIle 47
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 TTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAATC 341

Qy         48 LeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSer 67
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTCC 401

RESULT 6
US-09-747-155-366
; Sequence 366, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
;  PRIOR FILING DATE: 1999-12-22
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 366
;   LENGTH: 487
;   TYPE: DNA
;   ORGANISM: Saimiri boliviensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(487)
;   OTHER INFORMATION: Taxon = 27679; gene = SBO222; Accession DDBJ/EMBL/GenBank = AF1798
;   NAME/KEY: CDS
;   LOCATION: (2)..(487)
;   OTHER INFORMATION: Product = olfactory receptor
US-09-747-155-366

Alignment Scores:
Pred. No.:              3.66e-23        Length:         487
Score:                  33.00           Matches:        33
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            10.61%          Indels:         0
DB:                     9               Gaps:           0

US-10-023-597-24 (1-311) x US-09-747-155-366 (1-487)

Qy        172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGly 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAATGGC 205

Qy        192 SerTyrIleAsnValLeuValIlePheIleValValThr 204
              |||||||||||||||||||||||||||||||||||||||
Db        206 TCTTACATAAATGTTCTGGTCATCTTTATTGTTGTGACC 244

RESULT 7
US-09-747-155-232
; Sequence 232, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
;  PRIOR FILING DATE: 1999-12-22
```

```
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 232
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(485)
;   OTHER INFORMATION: Taxon = 9606; gene = HSA18; pseudogene; Accession DDBJ/EMBL/GenBa
US-09-747-155-232

Alignment Scores:
Pred. No.:              3.32e-10        Length:         485
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     9               Gaps:           0

US-10-023-597-24 (1-311) x US-09-747-155-232 (1-485)

Qy        171 AspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsn 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GACAACCTTGTCAATCACTATATGTGTGACATCCTTCCTCTCCTTGAGCTCTCCTGCAAC 201

RESULT 8
US-09-886-055-50
; Sequence 50, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
;  APPLICANT: STRYER, LUBERT
;  APPLICANT:  ZOZULYA, SERGEY
;  TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;  TITLE OF INVENTION:  BIOSENSORS OF CHEMICAL SENSANTS
;  FILE REFERENCE: 078003-0277150
;  CURRENT APPLICATION NUMBER: US/09/886,055
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 60/213,812
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 522
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-886-055-50

Alignment Scores:
Pred. No.:              6.36e-10        Length:         933
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     9               Gaps:           0

US-10-023-597-24 (1-311) x US-09-886-055-50 (1-933)

Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291

RESULT 9
US-09-804-291-50
; Sequence 50, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
;  APPLICANT: ZOZULA, SERGEY
;  TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
;  FILE REFERENCE: P 0278005
;  CURRENT APPLICATION NUMBER: US/09/804,291
;  CURRENT FILING DATE:  2001-03-13
;  PRIOR APPLICATION NUMBER: 60/188,914
;  PRIOR FILING DATE: 2000-03-13
;  PRIOR APPLICATION NUMBER: 60/192,033
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/198,474
;  PRIOR FILING DATE: 2000-04-14
;  PRIOR APPLICATION NUMBER: 60/199,335
;  PRIOR FILING DATE: 2000-04-24
;  PRIOR APPLICATION NUMBER: 60/207,702
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/213,849
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: 60/226,534
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: 60/230,732
;  PRIOR FILING DATE: 2000-09-07
;  PRIOR APPLICATION NUMBER: 60/266,862
;  PRIOR FILING DATE: 2001-02-07
;  NUMBER OF SEQ ID NOS: 529
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-804-291-50

Alignment Scores:
Pred. No.:              6.36e-10        Length:         933
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     10              Gaps:           0

US-10-023-597-24 (1-311) x US-09-804-291-50 (1-933)

Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291

RESULT 10
US-10-343-650A-395
; Sequence 395, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
;  APPLICANT: HAGA, TATSUYA
;  TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
;  FILE REFERENCE: 31671-186347
;  CURRENT APPLICATION NUMBER: US/10/343,650A
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: JP 2000/237818
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: JP 2001/34434
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 694
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 395
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(933)
US-10-343-650A-395

Alignment Scores:
Pred. No.:              6.36e-10        Length:         933
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     13              Gaps:           0

US-10-023-597-24 (1-311) x US-10-343-650A-395 (1-933)
```

```
Qy          79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291
```

RESULT 11
US-10-220-382-35
; Sequence 35, Application US/10220382
; Publication No. US20030119111A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: YAO, Monique G.
; APPLICANT: SHIH, Leo L.
; APPLICANT: TRIBOULEY, Catherine
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: YUE, Henry
; APPLICANT: KHAN, Farrah A.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: YANG, Junming
; APPLICANT: HARLAND, Lee
; APPLICANT: WALSH, Roderick T.
; APPLICANT: LO, Terence P.
; APPLICANT: BOROWSKY, Mark L.
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0044 PCT
; CURRENT APPLICATION NUMBER: US/10/220,382
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730
; PRIOR FILING DATE: 2000-03-03; 2000-03-10; 2000-03-17; 2000-03-20
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030119111A1 7472439CB1
US-10-220-382-35

Alignment Scores:

Pred. No.:	6.36e-10	Length:	933
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	15	Gaps:	0

US-10-023-597-24 (1-311) x US-10-220-382-35 (1-933)

```
Qy          79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291
```

RESULT 12
US-10-005-041A-13
; Sequence 13, Application US/10005041A
; Publication No. US20030232331A1
; GENERAL INFORMATION:
; APPLICANT: Casman, Stacie J
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gilbert, Jennifer A
; APPLICANT: Mayotte, Jane E
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Mishra, Vishnu
; APPLICANT: Vernet, Corine AM
; APPLICANT: Dickinson, Kevin S
; APPLICANT: Ballinger, Robert A
; APPLICANT: Wolenc, Adam R
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Gerlach, Valerie
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-215
; CURRENT APPLICATION NUMBER: US/10/005,041A
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/251,459
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 60/259,007
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 953
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-005-041A-13

Alignment Scores:

Pred. No.:	6.5e-10	Length:	953
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	16	Gaps:	0

US-10-023-597-24 (1-311) x US-10-005-041A-13 (1-953)

```
Qy          79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         247 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 306
```

RESULT 13
US-10-024-399-1
; Sequence 1, Application US/10024399
; Publication No. US20030100491A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Colman, Steven D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Ballinger, Robert A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Casman, Stacie J.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-224AE
; CURRENT APPLICATION NUMBER: US/10/024,399
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,635
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/259,743
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/299,327
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 60/261,498
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/263,689
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/267,464
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/271,021
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/275,946

; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/278,150
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/285,718
; PRIOR FILING DATE: 2001-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 976
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-024-399-1

Alignment Scores:

Pred. No.:	6.65e-10	Length:	976
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	15	Gaps:	0

US-10-023-597-24 (1-311) x US-10-024-399-1 (1-976)

```
Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 301
```

RESULT 14
US-10-017-161-325
; Sequence 325, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 325
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1333)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1133)
US-10-017-161-325

Alignment Scores:

Pred. No.:	9.06e-10	Length:	1333
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	15	Gaps:	0

US-10-023-597-24 (1-311) x US-10-017-161-325 (1-1333)

```
Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 491
```

RESULT 15
US-10-292-798-291
; Sequence 291, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 291
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1333)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1133)
US-10-292-798-291

Alignment Scores:

Pred. No.:	9.06e-10	Length:	1333
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	16	Gaps:	0

US-10-023-597-24 (1-311) x US-10-292-798-291 (1-1333)

```
Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 491
```

Search completed: September 30, 2004, 17:26:19
Job time : 550 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 08:16:15 ; Search time 4423 Seconds
(without alignments)
3047.634 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 1577
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163409_14879/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023597_@CGN_1_1_3731_@runat_29092004_163409_14879 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1577	100.0	963	6	AX555964	AX555964 Sequence
2	1568	99.4	994	6	AX241500	AX241500 Sequence
3	1568	99.4	172991	2	AC083958	AC083958 Homo sapi
c 4	1568	99.4	192926	9	AP000916	AP000916 Homo sapi
5	1333.5	84.6	935	6	AX241493	AX241493 Sequence
6	1309	83.0	936	10	AY073678	AY073678 Mus muscu
7	1309	83.0	936	10	AY318059	AY318059 Mus muscu
c 8	1309	83.0	201304	10	AC096785	AC096785 Mus muscu
c 9	1303	82.6	258873	2	AC096935	AC096935 Rattus no
10	1273.5	80.8	930	6	AX242246	AX242246 Sequence
11	1273.5	80.8	933	6	AX244609	AX244609 Sequence
12	1273.5	80.8	933	6	AX448443	AX448443 Sequence
13	1273.5	80.8	933	6	BD144473	BD144473 Novel G-p
14	1273.5	80.8	991	6	AX702842	AX702842 Sequence
15	1273.5	80.8	1333	6	AX646099	AX646099 Sequence
16	1273.5	80.8	1333	9	AB065834	AB065834 Homo sapi
c 17	1273.5	80.8	150847	9	AP001524	AP001524 Homo sapi
c 18	1273.5	80.8	172991	2	AC083958	AC083958 Homo sapi
19	1268	80.4	976	6	AX556206	AX556206 Sequence
20	1266.5	80.3	953	6	AX554459	AX554459 Sequence
21	1232.5	78.2	933	10	AY318057	AY318057 Mus muscu
22	1230.5	78.0	933	10	AY073346	AY073346 Mus muscu
23	1230.5	78.0	235933	2	AC131161	AC131161 Rattus no
24	1224.5	77.6	253120	2	AC097099	AC097099 Rattus no
25	1223.5	77.6	933	10	AY318056	AY318056 Mus muscu
26	1223.5	77.6	209420	2	AC109178	AC109178 Mus muscu
27	1221	77.4	931	6	AX646877	AX646877 Sequence
28	1220.5	77.4	933	10	AF282279	AF282279 Mus muscu
29	1220.5	77.4	933	10	AY073780	AY073780 Mus muscu
30	1220.5	77.4	933	10	AY318060	AY318060 Mus muscu
c 31	1220.5	77.4	209644	10	AC074314	AC074314 433m13, c
32	1220	77.4	1655	6	AX357037	AX357037 Sequence
c 33	1217.5	77.2	235933	2	AC131161	AC131161 Rattus no
34	1212.5	76.9	933	10	AY073208	AY073208 Mus muscu
35	1201.5	76.2	221405	2	AC114252	AC114252 Rattus no
36	1189.5	75.4	933	10	AF282282	AF282282 Mus muscu
37	1189.5	75.4	933	10	AY073207	AY073207 Mus muscu
38	1185.5	75.2	933	10	AY318058	AY318058 Mus muscu
39	1180	74.8	933	6	AX242071	AX242071 Sequence
40	1180	74.8	936	6	AX244610	AX244610 Sequence
41	1180	74.8	936	6	AX454190	AX454190 Sequence
42	1180	74.8	936	9	AF238488	AF238488 Homo sapi
43	1180	74.8	942	6	BD144474	BD144474 Novel G-p
44	1180	74.8	951	6	AX448445	AX448445 Sequence
45	1180	74.8	989	6	AX702946	AX702946 Sequence

ALIGNMENTS

RESULT 1

```
AX555964
LOCUS       AX555964                963 bp    DNA     linear   PAT 27-NOV-2002
DEFINITION  Sequence 23 from Patent WO0250275.
ACCESSION   AX555964
VERSION     AX555964.1  GI:25899399
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Padigaru,M., Kekuda,R., Li,L., Ballinger,R.A., Casman,S.J.,
            Spytek,K.A., Baumgartner,J.C. and Burgess,C.E.
  TITLE     Novel proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0250275-A 23 27-JUN-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .963
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.68e-111      Length:          963
Score:                    1577.00        Matches:         311
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-10-023-597-24 (1-311) x AX555964 (1-963)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 61

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 121

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 181

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 241

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 301

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAG 361

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 421

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 481

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 541

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 601

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 661

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 721

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 781

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 841

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 901

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        902 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 934

RESULT 2
AX241500
LOCUS       AX241500                994 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 248 from Patent WO0127158.
ACCESSION   AX241500
VERSION     AX241500.1  GI:15798375
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 248 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .994
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="(H38g97 nucleotide)"
ORIGIN

Alignment Scores:
Pred. No.:                8.4e-111       Length:          994
Score:                    1568.00        Matches:         310
Percent Similarity:       99.68%         Conservative:    0
Best Local Similarity:    99.68%         Mismatches:      1
Query Match:              99.43%         Indels:          0
DB:                       6              Gaps:            0

US-10-023-597-24 (1-311) x AX241500 (1-994)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 240
```

```
Qy          81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 300

Qy         101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTAG 360

Qy         121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 420

Qy         141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTGGGGCTGTGGCTCATACA 480

Qy         161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 540

Qy         181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 600

Qy         201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 660

Qy         221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 720

Qy         241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 780

Qy         261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 840

Qy         281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 900

Qy         301 LeuLysArgThrPheSerArgIleSerPheSer 311
               |||||||||||||||||||||||||||||||
Db         901 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 933
```

```
RESULT 3
AC083958
LOCUS       AC083958               172991 bp    DNA     linear   HTG 03-MAR-2001
DEFINITION  Homo sapiens clone RP11-18M5, WORKING DRAFT SEQUENCE, 8 unordered
            pieces.
ACCESSION   AC083958
VERSION     AC083958.2  GI:13184080
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 172991)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome, clone RP11-18M5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172991)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
            FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
            Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
            Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G.,
            Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
            O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
            Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 2, 2001 this sequence version replaced gi:10717225.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L11149
                Center clone name: 18_M_5
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 169811 bases at least Q40
                Consensus quality: 170912 bases at least Q30
                Consensus quality: 171602 bases at least Q20
                Insert size: 176000; agarose-fp
                Insert size: 172291; sum-of-contigs
                Quality coverage: 7.3 in Q20 bases; agarose-fp
                Quality coverage: 7.4 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      388: contig of 388 bp in length
            *      389      488: gap of 100 bp
            *      489     3383: contig of 2895 bp in length
            *     3384     3483: gap of 100 bp
            *     3484     6775: contig of 3292 bp in length
            *     6776     6875: gap of 100 bp
            *     6876    69914: contig of 63039 bp in length
            *    69915    70014: gap of 100 bp
            *    70015    87665: contig of 17651 bp in length
            *    87666    87765: gap of 100 bp
            *    87766   112075: contig of 24310 bp in length
            *   112076   112175: gap of 100 bp
            *   112176   166387: contig of 54212 bp in length
            *   166388   166487: gap of 100 bp
            *   166488   172991: contig of 6504 bp in length.
FEATURES             Location/Qualifiers
     source          1. .172991
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-18M5"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .388
                     /note="assembly_fragment
                     clone_end:SP6
```

```
                     vector_side:left"
     misc_feature    489. .3383
                     /note="assembly_fragment"
     misc_feature    3484. .6775
                     /note="assembly_fragment"
     misc_feature    6876. .69914
                     /note="assembly_fragment"
     misc_feature    70015. .87665
                     /note="assembly_fragment"
     misc_feature    87766. .112075
                     /note="assembly_fragment"
     misc_feature    112176. .166387
                     /note="assembly_fragment"
     misc_feature    166488. .172991
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:                1.68e-108      Length:          172991
Score:                    1568.00        Matches:         310
Percent Similarity:       99.68%         Conservative:    0
Best Local Similarity:    99.68%         Mismatches:      1
Query Match:              99.43%         Indels:          0
DB:                       2              Gaps:            0

US-10-023-597-24 (1-311) x AC083958 (1-172991)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94318 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 94377

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94378 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 94437

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94438 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 94497

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94498 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 94557

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94558 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 94617

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94618 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTAG 94677

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94678 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 94737

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94738 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 94797

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94798 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 94857

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94858 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 94917

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94918 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 94977

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94978 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 95037

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95038 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 95097

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95098 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 95157

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95158 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 95217

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db      95218 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 95250

RESULT 4
AP000916/c
LOCUS       AP000916              192926 bp    DNA     linear   PRI 01-FEB-2003
DEFINITION  Homo sapiens genomic DNA, chromosome 11 clone:RP11-728D14, complete
            sequence.
ACCESSION   AP000916
VERSION     AP000916.6  GI:28189528
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA
  JOURNAL   Published Only in Database (1999)
REFERENCE   2  (bases 1 to 192926)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     On Jan 31, 2003 this sequence version replaced gi:15637083.
FEATURES             Location/Qualifiers
     source          1. .192926
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q"
                     /clone="RP11-728D14"
ORIGIN

Alignment Scores:
Pred. No.:                1.88e-108      Length:          192926
Score:                    1568.00        Matches:         310
Percent Similarity:       99.68%         Conservative:    0
Best Local Similarity:    99.68%         Mismatches:      1
Query Match:              99.43%         Indels:          0
DB:                       9              Gaps:            0

US-10-023-597-24 (1-311) x AP000916 (1-192926)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102700 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 102641

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     102640 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 102581

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db     102580 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 102521

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db     102520 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 102461

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db     102460 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 102401

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db     102400 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTAG 102341

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db     102340 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 102281

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db     102280 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 102221

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db     102220 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 102161

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db     102160 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 102101

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db     102100 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 102041

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db     102040 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 101981

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db     101980 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 101921

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db     101920 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 101861

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db     101860 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 101801

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db     101800 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 101768
```

```
RESULT 5
AX241493
LOCUS       AX241493                935 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 241 from Patent WO0127158.
ACCESSION   AX241493
VERSION     AX241493.1  GI:15798368
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 241 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="(H38g90 nucleotide)"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                5.79e-93        Length:          935
Score:                    1333.50         Matches:         257
Percent Similarity:       92.18%          Conservative:    26
Best Local Similarity:    83.71%          Mismatches:      23
Query Match:              84.56%          Indels:          2
DB:                       6               Gaps:            1

US-10-023-597-24 (1-311) x AX241493 (1-935)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db          1 ATGCCTGCCAAGAACTCCTCCTCCGTGACAGCGTTTATCCTCTCAGGCTTAACCGACCAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db         61 CCGGGACTCCAGATCCCCGCCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        121 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db        181 TTTTTCCCCTTCAACTTGTCCTTCATAGATTTTAGTTATTCCACTACCCTCGCCCCTAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        241 ATGCTGATGAGCTTTGTCTCA---GAGAACATCATTTCCTATGCAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        298 CTTTTTTTCTTCTGTTTCTTTGTCTTTTCTGAATCCTATATTCTATCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        358 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGGTCACCATGTCTCCCCAGATG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        418 TGTTTGCTCCTTTTACTGGGTGTCTATGGGATGGGGAT-TTTGGGGCTGTGGCTCATATG 476

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        477 GGAAACATAATGTTTATGTCCTTTTGTGGAGACAACCTTGTCAATCACTATATGTGTGAC 536

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        537 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATAAATTTGCTGGTGGTTTTT 596

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        597 ATTATTGTGACCGTTGGCATTGGGGTGCCGATTGTCACCATTTTTCTCTCTTATGGTTTT 656

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        657 ATTCTTTCCAGCATTCTCCACATTAGTTCCACAGAGGGCAGGTCTAAAGCCTTCAGTACC 716

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        717 TGCAGTTCCCACATAATTGTGGTATCGCTTTTCTTTGGGTCAGGTGCTTTCATGTACCTC 776

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
```

```
Db        777 AAACCACCTTCTATTCTACCCCTGGACCAGGGGAAAGTGTCCTCCATTTTTTGTACTGCT 836

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        837 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAAGATGTCAAAGTTGCC 896

Qy        301 LeuLysArgThrPheSerArg 307
              |||:::|||||||||   |||
Db        897 CTGAGGAGAACCTTTTGCAGA 917

RESULT 6
AY073678
LOCUS       AY073678                936 bp    DNA     linear   ROD 03-APR-2002
DEFINITION  Mus musculus olfactory receptor MOR161-5 gene, complete cds.
ACCESSION   AY073678
VERSION     AY073678.1  GI:18480653
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 936)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)
  MEDLINE   21676863
   PUBMED   11802173
REFERENCE   2  (bases 1 to 936)
  AUTHORS   Young,J.M., Friedman,C., Williams,E.M., Ross,J.A., Tonnes-Priddy,L.
            and Trask,B.J.
  TITLE     Different evolutionary processes shaped the mouse and human
            olfactory receptor gene families
  JOURNAL   Hum. Mol. Genet. 11 (5), 535-546 (2002)
  MEDLINE   21864068
   PUBMED   11875048
REFERENCE   3  (bases 1 to 936)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .936
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="9A5"
                     /note="GA_x5J8B7W60AJ-1756545-1755610"
     mRNA            <1. .>936
                     /product="olfactory receptor MOR161-5"
     CDS             1. .936
                     /codon_start=1
                     /evidence=not_experimental
                     /product="olfactory receptor MOR161-5"
                     /protein_id="AAL61341.1"
                     /db_xref="GI:18480654"
                     /translation="MAAANTSSVAEFILVGLTDQPQLQIPLFFLFLGFYIVTMVGNLG
                     LITLIGLNSHLHIPMYFFLFNLSFIDFSYSTTLTPKMLVGFVLRKNIISYAGCMTQFF
                     FFCFFVFSESYILSAMAYDRYVAICKPLLYSVTMSPQVCSYLLSGVYGMGVFGAVAHM
                     GNLQFISFCADNIINHYMCDIIPLLELSCNSSYINLLVVFIVVTIGIGVPIVTIFISY
                     GFILSSILHISSKEGRSKAFSTCTSHIIVVSLFFGSGAFMYLKPPSSLPLDQGKVSSV
                     FYTAVVPMFNPLIYSLRNKDVKIALKKTLSRKNFS"
ORIGIN

Alignment Scores:
Pred. No.:              4.26e-91        Length:         936
Score:                  1309.00         Matches:        246
Percent Similarity:     89.07%          Conservative:   31
Best Local Similarity:  79.10%          Mismatches:     34
Query Match:            83.01%          Indels:         0
DB:                     10              Gaps:           0

US-10-023-597-24 (1-311) x AY073678 (1-936)
```

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||   ||:::|||||||||   ||||||||||||   ||||||      |||
Db          1 ATGGCTGCAGCGAACACCTCCTCAGTGGCAGAGTTCATCCTTGTTGGATTAACAGACCAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   ||||||:::|||:::||||||||||||||||||||||||   ||||||:::|||
Db         61 CCTCAACTCCAGATCCCACTCTTCTTCCTGTTTCTAGGTTTCTACATTGTCACCATGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||||||||||||||
Db        121 GGGAACCTGGGCTTGATCACACTGATAGGGCTGAACTCTCACCTGCATATCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||   :::||||||||:::|||||||||:::   ||||||
Db        181 TTTTTCCTCTTCAATTTGTCCTTCATAGATTTCAGTTACTCCACCACACTCACCCCTAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||:::   ||||||   ||||||||||||||||||:::   ||||||||||:::|||
Db        241 ATGCTGGTGGGGTTTGTTCTGAGGAAGAACATCATCTCCTATGCAGGGTGTATGACACAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        301 TTTTTCTTCTTCTGCTTCTTTGTCTTTTCTGAGTCTTACATCCTGTCAGCAATGGCATAT 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||   ||||||||||:::::|||||||||||||||||||
Db        361 GACCGCTATGTTGCCATCTGTAAACCACTGTTGTATAGTGTCACCATGTCTCCCCAAGTG 420

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||   ||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTCATACCTTTTGTCAGGTGTCTATGGGATGGGAGTTTTTGGGGCTGTGGCCCATATG 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||:::   |||:::::|||||||||||||||:::::|||||||||||||||||||
Db        481 GGAAACCTACAGTTTATAAGCTTCTGTGCTGACAACATCATTAACCACTATATGTGTGAC 540

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||:::|||||||||||||||||||||||||   ||||||||||:::|||||||:::|||
Db        541 ATAATTCCCCTCCTTGAGCTGTCCTGCAATAGCTCATACATCAATTTGTTGGTGGTTTTT 600

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||:::|||||||||||||||||||||   :::|||||||||||||||||
Db        601 ATTGTTGTGACCATTGGCATTGGAGTGCCCATTGTTACCATTTTTATATCTTATGGTTTC 660

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   :::|||||||   ||||||||||||||||||||||:::
Db        661 ATCCTCTCTAGCATTCTCCACATTAGTTCTAAAGAAGGCAGGTCGAAAGCCTTTAGTACA 720

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||:::|||:::|||||||   ||||||||||||||||||||||||||||||   ||||||
Db        721 TGTACGTCCCATATTATTGTAGTGTCTCTCTTTTTTGGGTCAGGAGCTTTTATGTACCTC 780

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||   ||||||||||||||||||||||||||||:::||||||||||
Db        781 AAACCACCTTCTAGTTTGCCTCTGGATCAGGGAAAAGTATCCTCTGTTTTCTACACCGCT 840

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        841 GTGGTGCCCATGTTCAATCCATTAATCTATAGCCTGAGGAATAAGGATGTCAAAATTGCT 900

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||:::|||   ||||||   :::||||||
Db        901 CTGAAGAAAACCTTGAGCAGAAAAAATTTCTCT 933

RESULT 7
AY318059
LOCUS       AY318059                936 bp    DNA     linear   ROD 14-OCT-2003
DEFINITION  Mus musculus olfactory receptor GA_x6K02T2PVTD-31540342-31541277
            (GA_x6K02T2PVTD-31540342-31541277) gene, complete cds.
ACCESSION   AY318059
VERSION     AY318059.1  GI:32063397
KEYWORDS    .
```

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 936)
  AUTHORS   Young,J.M., Shykind,B.M., Lane,R.P., Tonnes-Priddy,L., Ross,J.A.,
            Walker,M., Williams,E.M. and Trask,B.J.
  TITLE     Odorant receptor expressed sequence tags demonstrate olfactory
            expression of over 400 genes, extensive alternate splicing and
            unequal expression levels
  JOURNAL   Genome Biol. 4 (11), R71 (2003)
REFERENCE   2  (bases 1 to 936)
  AUTHORS   Sanders,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2003) Celera Discovery System, Celera Genomics,
            45 West Gude Drive, Rockville, MD 20850, USA
COMMENT     AY317244-AY318733 were obtained from Celera Discovery System Mouse
            Genome Assembly Release 13.  Some of these may be redundant with
            the previous dataset (AY072961-AY074256, AY096300-AY096356 and
            AF500625-AF500776) which was obtained from an earlier assembly
            release.
FEATURES             Location/Qualifiers
     source          1. .936
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="9A5"
     gene            <1. .>936
                     /gene="GA_x6K02T2PVTD-31540342-31541277"
                     /note="synonyms: MOR161-5, GA_x5J8B7W60AJ-1756545-1755610"
     mRNA            <1. .>936
                     /gene="GA_x6K02T2PVTD-31540342-31541277"
                     /product="olfactory receptor
                     GA_x6K02T2PVTD-31540342-31541277"
     CDS             1. .936
                     /gene="GA_x6K02T2PVTD-31540342-31541277"
                     /codon_start=1
                     /product="olfactory receptor
                     GA_x6K02T2PVTD-31540342-31541277"
                     /protein_id="AAP71356.1"
                     /db_xref="GI:32063398"
                     /translation="MAAANTSSVAEFILVGLTDQPQLQIPLFFLFLGFYIVTMVGNLG
                     LITLIGLNSHLHIPMYFFLFNLSFIDFSYSTTLTPKMLVGFVLRKNIISYAGCMTQFF
                     FFCFFVFSESYILSAMAYDRYVAICKPLLYSVTMSPQVCSYLLSGVYGMGVFGAVAHM
                     GNLQFISFCADNIINHYMCDIIPLLELSCNSSYINLLVVFIVVTIGIGVPIVTIFISY
                     GFILSSILHISSKEGRSKAFSTCTSHIIVVSLFFGSGAFMYLKPPSSLPLDQGKVSSV
                     FYTAVVPMFNPLIYSLRNKDVKIALKKTLSRKNFS"
ORIGIN

Alignment Scores:
Pred. No.:                4.26e-91        Length:          936
Score:                    1309.00         Matches:         246
Percent Similarity:       89.07%          Conservative:    31
Best Local Similarity:    79.10%          Mismatches:      34
Query Match:              83.01%          Indels:          0
DB:                       10              Gaps:            0

US-10-023-597-24 (1-311) x AY318059 (1-936)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||   |||:::|||||||||   ||||||||||||   ||||||       |||
Db          1 ATGGCTGCAGCGAACACCTCCTCAGTGGCAGAGTTCATCCTTGTTGGATTAACAGACCAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   ||||||:::|||:::||||||||||||||||||||||||   ||||||:::|||
Db         61 CCTCAACTCCAGATCCCACTCTTCTTCCTGTTTCTAGGTTTCTACATTGTCACCATGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||   ||||||||||||||||||||||
Db        121 GGGAAACCTGGGCTTGATCACACTGATAGGGCTGAACTCTCACCTGCATATCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||   :::|||||||||:::|||||||||:::   ||||||
Db        181 TTTTTCCTCTTCAATTTGTCCTTCATAGATTTCAGTTACTCCACCACACTCACCCCTAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||:::   ||||||   ||||||||||||||||||:::   ||||||||||:::|||
Db        241 ATGCTGGTGGGGTTTGTTCTGAGGAAGAACATCATCTCCTATGCAGGGTGTATGACACAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        301 TTTTTCTTCTTCTGCTTCTTTGTCTTTTCTGAGTCTTACATCCTGTCAGCAATGGCATAT 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||   ||||||||||||:::::|||||||||||||||||||
Db        361 GACCGCTATGTTGCCATCTGTAAACCACTGTTGTATAGTGTCACCATGTCTCCCCAAGTG 420

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||   ||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTCATACCTTTTGTCAGGTGTCTATGGGATGGGAGTTTTTGGGGCTGTGGCCCATATG 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||:::   |||::::::|||||||||||||||:::::|||||||||||||||||||
Db        481 GGAAACCTACAGTTTATAAGCTTCTGTGCTGACAACATCATTAACCACTATATGTGTGAC 540

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||:::|||||||||||||||||||||   ||||||||||:::|||||||:::|||
Db        541 ATAATTCCCCTCCTTGAGCTGTCCTGCAATAGCTCATACATCAATTTGTTGGTGGTTTTT 600

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||:::|||||||||||||||||||   :::|||||||||||||||||||
Db        601 ATTGTTGTGACCATTGGCATTGGAGTGCCCATTGTTACCATTTTTATATCTTATGGTTTC 660

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   :::|||||||   ||||||||||||||||||||||:::
Db        661 ATCCTCTCTAGCATTCTCCACATTAGTTCTAAAGAAGGCAGGTCGAAAGCCTTTAGTACA 720

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||:::|||:::|||||||   ||||||||||||||||||||||||||||||   ||||||
Db        721 TGTACGTCCCATATTATTGTAGTGTCTCTCTTTTTTGGGTCAGGAGCTTTTATGTACCTC 780

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||   ||||||||||||||||||||||||||||||:::|||||||||
Db        781 AAACCACCTTCTAGTTTGCCTCTGGATCAGGGAAAAGTATCCTCTGTTTTCTACACCGCT 840

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        841 GTGGTGCCCATGTTCAATCCATTAATCTATAGCCTGAGGAATAAGGATGTCAAAATTGCT 900

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              ||||||:::|||   ||||||   :::|||||||
Db        901 CTGAAGAAAACCTTGAGCAGAAAAAATTTCTCT 933

RESULT 8
AC096785/c
LOCUS       AC096785               201304 bp    DNA     linear   ROD 29-OCT-2003
DEFINITION  Mus musculus clone rp23-382i2 map 18 strain C57BL/6J, complete
            sequence.
ACCESSION   AC096785
VERSION     AC096785.29  GI:38016067
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 201304)
  AUTHORS   Jiang,X., Song,L. and Roe,B.A.
  TITLE     Mus musculus BAC Clone rp23-382i2
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 201304)
  AUTHORS   Jiang,X., Song,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-SEP-2001) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
```

```
REFERENCE   3  (bases 1 to 201304)
  AUTHORS   Jiang,X., Song,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   4  (bases 1 to 201304)
  AUTHORS   Jiang,X., Song,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   5  (bases 1 to 201304)
  AUTHORS   Jiang,X., Song,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Oct 29, 2003 this sequence version replaced gi:34365904.
            -------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------------------.
FEATURES             Location/Qualifiers
     source          1. .201304
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /map="18"
                     /clone="rp23-382i2"
                     /clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"
ORIGIN

Alignment Scores:
Pred. No.:                1.06e-88      Length:          201304
Score:                    1309.00       Matches:         246
Percent Similarity:       89.07%        Conservative:    31
Best Local Similarity:    79.10%        Mismatches:      34
Query Match:              83.01%        Indels:          0
DB:                       10            Gaps:            0

US-10-023-597-24 (1-311) x AC096785 (1-201304)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||   |||:::|||||||||   ||||||||||   ||||||      |||
Db     121967 ATGGCTGCAGCGAACACCTCCTCAGTGGCAGAGTTCATCCTTGTTGGATTAACAGACCAG 121908

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   ||||||:::|||:::|||||||||||||||||||||||||   ||||||:::|||
Db     121907 CCTCAACTCCAGATCCCACTCTTCTTCCTGTTTCTAGGTTTCTACATTGTCACCATGGTG 121848

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||   ||||||||||||||||||||||
Db     121847 GGGAACCTGGGCTTGATCACACTGATAGGGCTGAACTCTCACCTGCATATCCCCATGTAC 121788

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||   :::||||||||||:::||||||||||:::   ||||||
Db     121787 TTTTTCCTCTTCAATTTGTCCTTCATAGATTTCAGTTACTCCACCACACTCACCCCTAAA 121728

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||:::   ||||||   ||||||||||||||||||:::   ||||||||||:::|||
Db     121727 ATGCTGGTGGGGTTTGTTCTGAGGAAGAACATCATCTCCTATGCAGGGTGTATGACACAG 121668

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db     121667 TTTTTCTTCTTCTGCTTCTTTGTCTTTTCTGAGTCTTACATCCTGTCAGCAATGGCATAT 121608

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||   ||||||||||||:::::||||||||||||||||||
Db     121607 GACCGCTATGTTGCCATCTGTAAACCACTGTTGTATAGTGTCACCATGTCTCCCCAAGTG 121548

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||      ||||||   ||||||||||||||||||||||||||||||||||||||||
Db     121547 TGTTCATACCTTTTGTCAGGTGTCTATGGGATGGGAGTTTTTGGGGCTGTGGCCCATATG 121488

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||:::   |||::::::||||||||||||||||:::::||||||||||||||||||
Db     121487 GGAAACCTACAGTTTATAAGCTTCTGTGCTGACAACATCATTAACCACTATATGTGTGAC 121428

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||:::|||||||||||||||||||||||||   ||||||||||||:::||||||:::|||
Db     121427 ATAATTCCCCTCCTTGAGCTGTCCTGCAATAGCTCATACATCAATTTGTTGGTGGTTTTT 121368

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||:::|||||||||||||||||||||   :::|||||||||||||||||||
Db     121367 ATTGTTGTGACCATTGGCATTGGAGTGCCCATTGTTACCATTTTTATATCTTATGGTTTC 121308

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   :::||||||   ||||||||||||||||||||||||:::
Db     121307 ATCCTCTCTAGCATTCTCCACATTAGTTCTAAAGAAGGCAGGTCGAAAGCCTTTAGTACA 121248

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||:::|||:::||||||   ||||||||||||||||||||||||||   ||||||
Db     121247 TGTACGTCCCATATTATTGTAGTGTCTCTCTTTTTTGGGTCAGGAGCTTTTATGTACCTC 121188

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||   ||||||||||||||||||||||||||||:::|||||||||
Db     121187 AAACCACCTTCTAGTTTGCCTCTGGATCAGGGAAAAGTATCCTCTGTTTTCTACACCGCT 121128

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db     121127 GTGGTGCCCATGTTCAATCCATTAATCTATAGCCTGAGGAATAAGGATGTCAAAATTGCT 121068

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              ||||||:::|||   ||||||   :::||||||
Db     121067 CTGAAGAAAACCTTGAGCAGAAAAAATTTCTCT 121035

RESULT 9
AC096935/c
LOCUS       AC096935              258873 bp    DNA     linear   HTG 10-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-154L7, *** SEQUENCING IN PROGRESS
            ***, 4 unordered pieces.
ACCESSION   AC096935
VERSION     AC096935.7  GI:30521491
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 258873)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
```

```
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 258873)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 258873)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:24953430.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GHMU
                Center clone name: CH230-154L7
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 207030 bases at least Q40
                Consensus quality: 209712 bases at least Q30
                Consensus quality: 211447 bases at least Q20
                Estimated insert size: 223149; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   245424: contig of 245424 bp in length
            *  245425   245524: gap of unknown length
            *  245525   246877: contig of 1353 bp in length
            *  246878   246977: gap of unknown length
            *  246978   254737: contig of 7760 bp in length
            *  254738   254837: gap of unknown length
            *  254838   258873: contig of 4036 bp in length.
FEATURES            Location/Qualifiers
     source         1. .258873
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-154L7"
     misc_feature   501. .1322
                    /note="clone_boundary
                    clone_end:T7
                    site:EcoRI
                    end_sequence:BZ107537"
     misc_feature   243527. .245424
                    /note="wgs_contig"
ORIGIN

Alignment Scores:
Pred. No.:              3.92e-88        Length:         258873
Score:                  1303.00         Matches:        245
Percent Similarity:     88.10%          Conservative:   29
Best Local Similarity:  78.78%          Mismatches:     37
Query Match:            82.63%          Indels:         0
DB:                     2               Gaps:           0

US-10-023-597-24 (1-311) x AC096935 (1-258873)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||  |||||||||:::|||||||||  ||||||||||||  ||||||       |||
Db     155892 ATGACTGCAGAGAACACCTCCTCGGTGGCAGAGTTCATCCTTGTTGGATTAACAGACCAG 155833

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||  |||||:::|||:::|||  ||||||||||||||||||  ||||||:::|||
Db     155832 CCTGACCTCCAGATCCCACTCTTCATCCTCTTTCTCGGTTTCTACATGGTCACTCTGGTG 155773

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||  ||||||||||||||||||  ||||||||||||||||||
Db     155772 GGGAACCTGGGCTTGATCACACTGATAGGGTTGAACTCTCACCTGCATATTCCCATGTAC 155713

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||      |||||||||  :::|||||||||:::|||||||||:::  ||||||
Db     155712 TTTTTCCTCCTCAACTTGTCCTTCATAGATTTCAGTTACTCTACTACACTTACCCCTAAA 155653

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||:::  ||||||  ||||||||||||||||||:::  |||||||||:::|||
Db     155652 ATGCTGGTGGGCTTTGTTGTGAGGAAGAACATCATCTCCTACGCAGGGTGTATGACTCAG 155593

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db     155592 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTACATCTTGTCGGCAATGGCATAC 155533

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||  ||||||  ||||||||||  :::||||||||||||||||||
Db     155532 GACCGCTATGTTGCCATCTGTAAACCATTGTTGTATAGGGTCACCATGTCTCCCCAAGTG 155473

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||  ||||||  ||||||||||||||||||||||||||||||||||||||||||
Db     155472 TGCTCATTCCTTTTGTCAGGCGTCTATGGGATGGGGGTTTTTGGGGCTGTGGCCCATATG 155413

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||:::  |||:::|||||||||||||||||:::::|||||||||||||||||||
Db     155412 GGAAACTTACAGTTTATAACCTTCTGTGCTGACAACATCATTAACCACTATATGTGTGAC 155353
```

```
Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||:::|||||||||||||||||||||||||||   |||||||||||:::|||||||:::|||
Db     155352 ATCATTCCCCTCCTTGAGCTGTCCTGCAATAGCTCTTACATCAATTTGTTGGTGGTTTTT 155293

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||:::|||||||||||||||||||||||   :::||||||||||||||||||
Db     155292 ATTGTTGTAACCATTGGCATTGGAGTACCCATTGTTACCATTTTTATATCTTATGGCTTC 155233

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   :::||||||   |||||||||||||||||||||||:::
Db     155232 ATCCTCTCTAGCATTCTCCACATTAGTTCCAAGGAAGGCAGGTCCAAAGCCTTTAGTACC 155173

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||:::|||:::||||||   ||||||||||||||||||||||||||||   ||||||
Db     155172 TGCACTTCTCATATTATTGTAGTGTCTCTCTTTTTTGGGTCAGGAGCTTTTATGTATCTC 155113

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||   |||||||||||||||||||||||||||:::||||||||||
Db     155112 AAACCACCTTCTAGTTTGCCTCTGGATCAGGGAAAAGTGTCCTCTGTTTTCTATACTGCT 155053

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db     155052 GTAGTGCCCATGTTCAATCCATTAATCTATAGCCTGAGGAATAAGGATGTCAAAATTGCT 154993

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              ||||||:::|||   ||||||   :::|||||||
Db     154992 CTGAAGAAAACCTTGAGCAGAAAAAATTTCTCT 154960

RESULT 10
AX242246
LOCUS       AX242246                930 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 994 from Patent WO0127158.
ACCESSION   AX242246
VERSION     AX242246.1  GI:15799121
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 994 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="(H38g844 nucleotide)"
ORIGIN

Alignment Scores:
Pred. No.:                2.14e-88        Length:         930
Score:                    1273.50         Matches:        250
Percent Similarity:       87.46%          Conservative:   22
Best Local Similarity:    80.39%          Mismatches:     38
Query Match:              80.75%          Indels:         1
DB:                       6               Gaps:           1

US-10-023-597-24 (1-311) x AX242246 (1-930)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::|||:::|||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||   ||||||||||||   ||||||:::|||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||||:::||||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   |||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::||||||||||   ||||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   ||||||   |||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||||:::
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||   |||   ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 11
AX244609
LOCUS       AX244609                933 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 35 from Patent WO0166742.
ACCESSION   AX244609
VERSION     AX244609.1  GI:15859510
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lal,P., Tang,Y.T., Patterson,C., Yao,M.G., Shih,L.L.,
            Tribouley,C.M., Lu,D.A., Yue,H., Khan,F.A., Policky,J.L.,
            Au-Young,J., Yang,J., Harland,L., Walsh,R.T., Lo,T.P. and
            Borowsky,M.L.
  TITLE     G-protein coupled receptors
  JOURNAL   Patent: WO 0166742-A 35 13-SEP-2001;
            Incyte Genomics, Inc. (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472439CB1"
ORIGIN

Alignment Scores:
Pred. No.:              2.15e-88        Length:         933
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     6               Gaps:           1

US-10-023-597-24 (1-311) x AX244609 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::::|::|||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 ||||||||||||||||||   ||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||   ||||||||||      ||||||:::||||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||||:::||||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::||||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   ||||||   ||||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::|||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||      |||      ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 12
AX448443
LOCUS       AX448443                933 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 109 from Patent WO0224726.
ACCESSION   AX448443
VERSION     AX448443.1  GI:21697342
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Veithen,A.
  TITLE     Olfactory and pheromones g-protein coupled receptors
  JOURNAL   Patent: WO 0224726-A 109 28-MAR-2002;
            ChemCom S.A. (BE)
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .933
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD37543.1"
                     /db_xref="GI:21697343"
                     /db_xref="REMTREMBL:CAD37543"
                     /translation="MAAKNSSVTEFILEGLTHQPGLRIPLFFLFLGFYTVTVVGNLGL
                     ITLIGLNSHLHTPMYFFLFNLSLIDFCFSTTITPKMLMSFVSRKNIISFTGCMTQLFF
                     FCFFVVSESFILSAMAYDRYVAICNPLLYTVTMSCQVCLLLLLGAYGMGFAGAMAHTG
                     SIMNLTFCADNLVNHFMCDILPLLELSCNSSYMNELVVFIVVAVDVGMPIVTVFISYA
                     LILSSILHNSSTEGRSKAFSTCSSHIIVVSLFFGSGAFMYLKPLSILPLEQGKVSSLF
                     YTIIVPVLNPLIYSLRNKDVKVALRRTLGRKIFS"
ORIGIN

Alignment Scores:
Pred. No.:              2.15e-88        Length:         933
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     6               Gaps:           1

US-10-023-597-24 (1-311) x AX448443 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::::|::|||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297
```

```
Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||    ||||||||||||       ||||||:::||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||:::|||||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||    ||||||:::|||    ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||      ||||||   ||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||     |||      ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 13
BD144473
LOCUS       BD144473                933 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Novel G-protein coupled receptors.
ACCESSION   BD144473
VERSION     BD144473.1  GI:27850231
KEYWORDS    JP 2002112793-A/198.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   Haga,T., Takeda,S. and Miyake,N.
  TITLE     Novel G-protein coupled receptors
  JOURNAL   Patent: JP 2002112793-A 198 16-APR-2002;
            JAPAN SCIENCE AND TECHNOLOGY CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002112793-A/198
            PD   16-APR-2002
            PF   09-FEB-2001 JP 2001034434
            PI   TATSUYA HAGA,SHIGEKI TAKEDA,NARIKI MIYAKE
            PC
            C12N15/09,A01K67/027,A61K38/00,A61K39/395,A61K39/395,A61K45/00, PC
             A61K48/00,
            PC   A61P43/00,C07K14/705,C07K16/28,C07K19/00,C12N1/15,C12N1/19, PC
              C12N1/21,
            PC   C12N5/10,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC
            G01N33/566//
            PC   C12P21/08,C12N15/00,A61K37/02,C12N5/00
            CC   Novel G-protein coupled receptors
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(933).
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.15e-88        Length:         933
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     6               Gaps:           1

US-10-023-597-24 (1-311) x BD144473 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   ||||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||:::::::|||:|||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::|||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||    ||||||||||||       ||||||:::||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||:::|||||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||    ||||||:::|||    ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||      ||||||   ||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777
```

```
Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              |||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||      |||      ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 14
AX702842
LOCUS       AX702842                 991 bp    DNA     linear   PAT 03-APR-2003
DEFINITION  Sequence 71 from Patent WO02059313.
ACCESSION   AX702842
VERSION     AX702842.1  GI:29537953
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Li,L., Ballinger,R.A., Padigaru,M., Kekuda,R., Colman,S.D.,
            Spytek,K.A., Casman,S.J., Vernet,C.A., Shenoy,S.G., Gusev,V.,
            Malyankar,U.M., Edinger,S., Gerlach,V., Smithson,G., Stone,D.J.,
            Sciore,P., Macdougall,J.R., Gunther,E., Peyman,J.A., Ellerman,K.,
            Gangolli,E.A. and Millet,I.
  TITLE     G-protein coupled receptors and nucleic acids encoding same
  JOURNAL   Patent: WO 02059313-A 71 01-AUG-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .991
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.29e-88        Length:         991
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     6               Gaps:           1

US-10-023-597-24 (1-311) x AX702842 (1-991)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||   ||||||   ||||||
Db         49 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 105

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::::|||:::|||||||||||||||||||||||   ||||||||||||
Db        106 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 165

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||   ||||||   |||||||||
Db        166 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 225

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        226 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 285

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        286 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 345

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||   ||||||||||||||||||||||||||
Db        346 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 405

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||||:::|||||||||   ||||||
Db        406 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 465

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||   ||||||||||||   |||||||:::|||||||||
Db        466 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 525

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||||:::|||||||||
Db        526 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 585

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        586 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 645

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   |||||||||||||||
Db        646 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 705

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   ||||||   ||||||||||||||||||||||||:::
Db        706 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 765

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||||   ||||||
Db        766 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 825

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||||||
Db        826 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 885

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||:::|||
Db        886 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 945

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||      |||      ||||||
Db        946 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 978

RESULT 15
AX646099
LOCUS       AX646099                1333 bp    DNA     linear   PAT 04-MAR-2003
DEFINITION  Sequence 291 from Patent EP1270724.
ACCESSION   AX646099
VERSION     AX646099.1  GI:28798478
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.
  TITLE     Guanosine triphosphate-binding protein coupled receptors
  JOURNAL   Patent: EP 1270724-A 291 02-JAN-2003;
            National Institute of Advanced Industrial Science and Technology
            (JP) ; Center for Advanced Science and Technology Incubation, Ltd.
            (JP)
FEATURES             Location/Qualifiers
     source          1. .1333
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             201. .1133
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD69184.1"
                     /db_xref="GI:28798479"
                     /db_xref="REMTREMBL:CAD69184"
                     /translation="MAAKNSSVTEFILEGLTHQPGLRIPLFFLFLGFYTVTVVGNLGL
                     ITLIGLNSHLHTPMYFFLFNLSLIDFCFSTTITPKMLMSFVSRKNIISFTGCMTQLFF
                     FCFFVVSESFILSAMAYDRYVAICNPLLYTVTMSCQVCLLLLLGAYGMGFAGAMAHTG
                     SIMNLTFCADNLVNHFMCDILPLLELSCNSSYMNELVVFIVVAVDVGMPIVTVFISYA
```

LILSSILHNSSTEGRSKAFSTCSSHIIVVSLFFGSGAFMYLKPLSILPLEQGKVSSLF
YTIIVPVLNPLIYSLRNKDVKVALRRTLGRKIFS"

ORIGIN

Alignment Scores:

Pred. No.:	3.1e-88	Length:	1333
Score:	1273.50	Matches:	250
Percent Similarity:	87.46%	Conservative:	22
Best Local Similarity:	80.39%	Mismatches:	38
Query Match:	80.75%	Indels:	1
DB:	6	Gaps:	1

US-10-023-597-24 (1-311) x AX646099 (1-1333)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||   ||||||   ||||||
Db        201 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 257

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||::::::|||:::|||||||||||||||||||||||   ||||||||||||
Db        258 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 317

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||   ||||||   |||||||||
Db        318 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 377

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        378 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 437

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        438 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 497

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                |||||||||||||||||||   ||||||||||||||||||||||||||||
Db        498 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 557

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||:::|||||||||   ||||||
Db        558 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 617

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||   ||||||||||||   ||||||:::|||||||||
Db        618 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 677

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||:::|||||||||
Db        678 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 737

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        738 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 797

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||||
Db        798 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 857

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   ||||||   ||||||||||||||||||||||||:::
Db        858 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 917

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||:::||||||   ||||||||||||||||||||||||||||   ||||||
Db        918 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 977

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||
Db        978 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 1037

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||:::|||
Db       1038 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 1097
```

```
Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::|||||||   |||   ||||||
Db       1098 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 1130
```

Search completed: September 30, 2004, 11:26:33
Job time : 4638 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 08:14:39 ; Search time 447 Seconds
(without alignments)
2955.681 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 1577
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163409_14869/app_query.fasta_1.455
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023597_@CGN_1_1_470_@runat_29092004_163409_14869 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1577	100.0	963	6	AAF88413	Aaf88413 Human GPC
2	1568	99.4	994	4	AAH31675	Aah31675 Human olf
3	1515	96.1	960	9	ADD12748	Add12748 cDNA enco
4	1333.5	84.6	935	4	AAH31668	Aah31668 Human olf
5	1273.5	80.8	930	4	AAH32421	Aah32421 Human olf
6	1273.5	80.8	933	5	AAS15910	Aas15910 DNA encod
7	1273.5	80.8	933	5	AAS42233	Aas42233 Human cDN
8	1273.5	80.8	933	6	ABZ43067	Abz43067 Human GPC
9	1273.5	80.8	933	6	ABK68455	Abk68455 Human DNA
10	1273.5	80.8	933	6	ABK37521	Abk37521 DNA encod
11	1273.5	80.8	991	6	ABS58782	Abs58782 Human G-p
12	1273.5	80.8	1333	9	ADC85838	Adc85838 Human GPC
13	1268	80.4	976	6	ABK97207	Abk97207 Human G-p
14	1268	80.4	976	9	AAD60328	Aad60328 Human G-p
15	1266.5	80.3	953	6	ABT05655	Abt05655 GPCR 4 pr
16	1221	77.4	931	9	ADC86616	Adc86616 Human GPC
17	1220	77.4	1655	6	AAD24449	Aad24449 Human G-p
18	1180	74.8	933	4	AAH32246	Aah32246 Human olf
19	1180	74.8	936	5	AAS15911	Aas15911 DNA encod
20	1180	74.8	936	6	ABK40195	Abk40195 Human G p
21	1180	74.8	942	6	ABZ43068	Abz43068 Human GPC
22	1180	74.8	951	6	ABK68456	Abk68456 Human DNA
23	1180	74.8	989	6	ABS58834	Abs58834 Human G-p
24	1180	74.8	1336	9	ADC86200	Adc86200 Human GPC
25	1073	68.0	942	7	ABZ77877	Abz77877 Human G p
26	1063	67.4	952	9	ADD12784	Add12784 cDNA enco
27	1058	67.1	952	6	AAF88431	Aaf88431 Human GPC
28	1048	66.5	961	6	ABQ88087	Abq88087 Human GPC
29	1035	65.6	962	9	ADC79403	Adc79403 Human G p
30	1024.5	65.0	934	6	AAF88464	Aaf88464 Human GPC
31	1023.5	64.9	945	7	ABZ77909	Abz77909 Human G p
32	1023	64.9	981	6	AAF88442	Aaf88442 Human GPC
33	1018.5	64.6	934	9	ADD12850	Add12850 cDNA enco
34	1008.5	64.0	984	4	AAH32275	Aah32275 Human olf
35	1007	63.9	981	9	ADD12806	Add12806 cDNA enco
36	1006	63.8	950	6	AAF88445	Aaf88445 Human GPC
37	1006	63.8	953	6	AAF88433	Aaf88433 Human GPC
38	1001.5	63.5	937	6	ABS58757	Abs58757 Human G-p
39	1001.5	63.5	937	6	ABS58759	Abs58759 Human G-p
40	1001.5	63.5	937	6	ABS58758	Abs58758 Human G-p
41	1000	63.4	950	9	ADD12812	Add12812 cDNA enco
42	999.5	63.4	942	5	AAS42218	Aas42218 Human cDN
43	999.5	63.4	942	6	ABK37508	Abk37508 DNA encod
44	999.5	63.4	1069	3	AAA38958	Aaa38958 Human G p
45	999.5	63.4	1342	9	ADC85834	Adc85834 Human GPC

ALIGNMENTS

RESULT 1
AAF88413
ID AAF88413 standard; cDNA; 963 BP.
XX
AC AAF88413;
XX
DT 12-NOV-2002 (first entry)
XX
DE Human GPCRX cDNA SEQ ID 23.
XX
KW Human; anti-HIV; nootropic; antiasthmatic; antiarteriosclerotic; GPCRX;
KW immunosuppressive; immunomodulator; cytostatic; antiinflammatory; AIDS;
KW antidiabetic; neuroprotective; anorectic; haemostatic; antibacterial;
KW fungicide; protozoal; virucide; human G-protein coupled receptor;
KW gene therapy; vaccine; cardiomyopathy; atherosclerosis; diabetes;
KW cell signal processing; cancer; obesity; neurodegenerative disorder;
KW cachexia; anorexia; Alzheimer's disease; Parkinson's disease;
KW immune disorder; graft versus host disease; bronchial asthma;
KW Crohn's disease; multiple sclerosis; haemophilia; infectious disease;
KW idiopathic thrombocytopenic purpura; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200250275-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US048958.
XX
PR 18-DEC-2000; 2000US-0256635P.
PR 21-DEC-2000; 2000US-0257876P.
PR 04-JAN-2001; 2001US-0259743P.

PR 10-JAN-2001; 2001US-0260718P.
PR 12-JAN-2001; 2001US-0261498P.
PR 24-JAN-2001; 2001US-0263689P.
PR 08-FEB-2001; 2001US-0267464P.
PR 22-FEB-2001; 2001US-0271021P.
PR 14-MAR-2001; 2001US-0275946P.
PR 23-MAR-2001; 2001US-0278150P.
PR 18-APR-2001; 2001US-0284591P.
PR 23-APR-2001; 2001US-0285718P.
PR 19-JUN-2001; 2001US-0299327P.
PR 16-AUG-2001; 2001US-0312902P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Kekuda R, Li L, Ballinger RA, Casman SJ, Spytek KA;
PI Baumgartner JC, Burgess CE;
XX
DR WPI; 2002-657419/70.
DR P-PSDB; AAB71174.
XX
PT New G-protein coupled receptor polypeptides, useful in gene therapy,
PT particularly for treating or preventing cardiomyopathy, atherosclerosis,
PT diabetes, Crohn's disease, hemophilia or cancer in humans.
XX
PS Claim 3; Page 90; 155pp; English.
XX
CC This invention describes novel human G-protein coupled receptor (GPCRX)
CC polypeptides which have anti-HIV, nootropic, antiasthmatic, protozoal;
CC antiarteriosclerotic; immunosuppressive, immunomodulator, cytostatic,
CC antiinflammatory, antidiabetic, neuroprotective, anorectic, haemostatic,
CC antibacterial, fungicide and virucide activity. The products of the
CC invention can be used in gene therapy or for vaccines. The GPCRX
CC polypeptide, GPCRX nucleic acid and antibody are useful for treating,
CC preventing or alleviating a GPCRX-associated disorder or a pathological
CC state in a subject e.g cardiomyopathy, atherosclerosis, diabetes, or a
CC disorder related to cell signal processing and metabolic pathway
CC modulation. The GPCRX polypeptide and nucleic acid are also useful for
CC diagnosing the presence of or predisposition to a disease associated with
CC altered levels of GPCRX, particularly cancer. These polypeptides, nucleic
CC acids and antibodies are also useful for treating or preventing obesity,
CC neurodegenerative disorders, AIDS, cancer-associated cachexia, anorexia,
CC Alzheimer's disease, Parkinson's disease, immune disorders, graft versus
CC host disease, bronchial asthma, Crohn's disease, multiple sclerosis,
CC haemophilia, idiopathic thrombocytopenic purpura or infectious disease.
CC They can also be used to screen for potential agonist and antagonist
CC compounds. The polypeptides are also useful as immunogens to produce
CC antibodies or as vaccines. Anti-GPCRX antibodies can be used
CC diagnostically to monitor protein levels in tissue as part of a clinical
CC testing procedure such as in determining the efficacy of a given
CC treatment regimen. The host cells are useful in producing non-human
CC transgenic animals which are useful for studying the function and/or
CC activity of GPCRX protein and for identifying and/or evaluating
CC modulators of GPCRX protein activity. AAF88402-AAF88465 encode the human
CC GPCRX proteins represented in AAB71163-AAB71226
XX
SQ Sequence 963 BP; 193 A; 247 C; 204 G; 319 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.12e-137	Length:	963
Score:	1577.00	Matches:	311
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-023-597-24 (1-311) x AAF88413 (1-963)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 61

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 121

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 181

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 241

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 301

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAG 361

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 421

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 481

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 541

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 601

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 661

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 721

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 781

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 841

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 901

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        902 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 934
```

RESULT 2
AAH31675
ID AAH31675 standard; DNA; 994 BP.
XX
AC AAH31675;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 248.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.

XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027582.
XX
PR 08-OCT-1999; 99US-0158615P.
PR 24-FEB-2000; 2000US-0184809P.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists.
XX
PS Claim 8; Page 276; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides which
CC encode polypeptides involved in olfactory sensation. The polynucleotides
CC can be used in screening for olfactory agonists and antagonists. The
CC methods allow for the determination of primary scents and the
CC identification of the odour receptors used to detect these primary
CC scents. The methods also enable determination of secondary scents and the
CC identification of combinations of odour receptors that are involved in
CC detecting such secondary scents. This enables the construction of a scent
CC representation (also called a scent fingerprint or scent profile), which
CC may be used to re-create and edit scents. Libraries of olfactory
CC receptors are useful for determining the interaction pattern of a
CC composition with the receptors, and can be used for determining
CC differences in the olfactory faculties of different individuals
XX
SQ Sequence 994 BP; 201 A; 250 C; 207 G; 336 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.93e-136	Length:	994
Score:	1568.00	Matches:	310
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.43%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x AAH31675 (1-994)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTAG 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 420

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 540

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 600

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 660

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 720

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 780

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 840

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 900

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        901 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 933
```

RESULT 3
ADD12748
ID ADD12748 standard; cDNA; 960 BP.
XX
AC ADD12748;
XX
DT 01-JAN-2004 (first entry)
XX
DE cDNA encoding novel human olfactory receptor (OR) seq id 12.
XX
KW cardiant; antiarteriosclerotic; antidiabetic;
KW G-Protein Coupled Receptor modulator; gene therapy; olfactory receptor;
KW G-Protein Coupled Receptor X-associated disorder;
KW GPCRX associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; cancer; diabetes;
KW human; G-protein coupled receptor; GPCR; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003109692-A1.
XX
PD 12-JUN-2003.
XX
PF 18-DEC-2001; 2001US-00023597.
XX
PR 18-DEC-2000; 2000US-0256635P.
PR 21-DEC-2000; 2000US-0257876P.
PR 04-JAN-2001; 2001US-0259743P.
PR 10-JAN-2001; 2001US-0260718P.
PR 12-JAN-2001; 2001US-0261498P.
PR 24-JAN-2001; 2001US-0263689P.
PR 08-FEB-2001; 2001US-0267464P.
PR 22-FEB-2001; 2001US-0271021P.
PR 14-MAR-2001; 2001US-0275946P.

```
PR   23-MAR-2001; 2001US-0278150P.
PR   18-APR-2001; 2001US-0284591P.
PR   23-APR-2001; 2001US-0285718P.
PR   19-JUN-2001; 2001US-0299327P.
PR   16-AUG-2001; 2001US-0312902P.
XX
PA   (PADI/) PADIGARU M.
PA   (KEKU/) KEKUDA R.
PA   (LILL/) LI L.
PA   (BALL/) BALLINGER R A.
PA   (CASM/) CASMAN S J.
PA   (SPYT/) SPYTEK K A.
PA   (BAUM/) BAUMGARTNER J C.
PA   (BURG/) BURGESS C E.
XX
PI   Padigaru M, Kekuda R, Li L, Ballinger RA, Casman SJ, Spytek KA;
PI   Baumgartner JC, Burgess CE;
XX
DR   WPI; 2003-801280/75.
DR   P-PSDB; ADD12749.
XX
PT   New isolated olfactory receptor polypeptide for treating or preventing a
PT   G-Protein Coupled Receptor associated disorder e.g. cardiomyopathy or
PT   atherosclerosis.
XX
PS   Claim 8; SEQ ID NO 23; 70pp; English.
XX
CC   The invention describes an isolated olfactory receptor polypeptide (I).
CC   (I) is used to identify an agent that binds to it by contacting (I) with
CC   the agent and determining binding. The agent is a cellular receptor or a
CC   downstream effector. (I) Is also used to identify an agent that modulates
CC   the expression or activity of (I) by contacting a cell expressing (I)
CC   with the agent and determining whether the agent modulates expression or
CC   activity of (I). (I), or nucleic acid encoding (I), is used to treat or
CC   prevent a G-Protein Coupled Receptor (GPCR)X-associated disorder in a
CC   subject, preferably a human, such as cardiomyopathy or atherosclerosis,
CC   or a disorder related to cell signal processing or metabolic pathway
CC   modulation. (I) Is used to determine the presence or predisposition to a
CC   disease associated with altered levels of (I), preferably cancer, in a
CC   mammal. A nucleic acid encoding (I) is used to determine the presence or
CC   predisposition to a disease associated with altered levels of the nucleic
CC   acid, preferably cancer, in a mammalian subject. An antibody that binds
CC   to (I) is used to treat or prevent a GPCRX-associated disorder, such as
CC   diabetes or a disorder related to cell signal processing and metabolic
CC   pathway modulation. An antibody that specifically binds (I) is used to
CC   determine the presence of amount of (I) in a sample by contacting the
CC   sample with the antibody and determining the presence or amount of bound
CC   antibody. This sequence encodes a novel human olfactory receptor (OR), a
CC   G protein coupled receptor (GPCR).
XX
SQ   Sequence 960 BP; 193 A; 247 C; 202 G; 318 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.35e-131      Length:          960
Score:                    1515.00        Matches:         309
Percent Similarity:       99.36%         Conservative:    0
Best Local Similarity:    99.36%         Mismatches:      2
Query Match:              96.07%         Indels:          2
DB:                       9              Gaps:            0

US-10-023-597-24 (1-311) x ADD12748 (1-960)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 61

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 121

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 181

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 241

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 301

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAG 361

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 421

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGTT-TTTGGGGCTGTGGCTCATACA 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACA--TGTGAC 538

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 598

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 658

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 718

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 778

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 838

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 898

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        899 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 931

RESULT 4
AAH31668
ID   AAH31668 standard; DNA; 935 BP.
XX
AC   AAH31668;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polynucleotide, SEQ ID NO: 241.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200127158-A2.
XX
PD   19-APR-2001.
XX
```

PF 06-OCT-2000; 2000WO-US027582.
XX
PR 08-OCT-1999; 99US-0158615P.
PR 24-FEB-2000; 2000US-0184809P.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists.
XX
PS Claim 8; Page 273; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides which
CC encode polypeptides involved in olfactory sensation. The polynucleotides
CC can be used in screening for olfactory agonists and antagonists. The
CC methods allow for the determination of primary scents and the
CC identification of the odour receptors used to detect these primary
CC scents. The methods also enable determination of secondary scents and the
CC identification of combinations of odour receptors that are involved in
CC detecting such secondary scents. This enables the construction of a scent
CC representation (also called a scent fingerprint or scent profile), which
CC may be used to re-create and edit scents. Libraries of olfactory
CC receptors are useful for determining the interaction pattern of a
CC composition with the receptors, and can be used for determining
CC differences in the olfactory faculties of different individuals
XX
SQ Sequence 935 BP; 189 A; 238 C; 190 G; 318 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.63e-114	Length:	935
Score:	1333.50	Matches:	257
Percent Similarity:	92.18%	Conservative:	26
Best Local Similarity:	83.71%	Mismatches:	23
Query Match:	84.56%	Indels:	2
DB:	4	Gaps:	1

US-10-023-597-24 (1-311) x AAH31668 (1-935)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||   |||:::|||||||||||||||||||   |||||||||:::||||||      |||
Db          1 ATGCCTGCCAAGAACTCCTCCTCCGTGACAGCGTTTATCCTCTCAGGCTTAACCGACCAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||:::|||   ||||||||||||||||||||||||||||||||||||||||
Db         61 CCGGGACTCCAGATCCCCGCCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||    :::|||||||||:::|||||||||:::    ||||||
Db        181 TTTTTCCCCTTCAACTTGTCCTTCATAGATTTTAGTTATTCCACTACCCTCGCCCCTAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||    :::|||||||||||||:::    |||||||||:::|||
Db        241 ATGCTGATGAGCTTTGTCTCA---GAGAACATCATTTCCTATGCAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        298 CTTTTTTTCTTCTGTTTCTTTGTCTTTTCTGAATCCTATATTCTATCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||:::|||||||||||||||||:::
Db        358 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGGTCACCATGTCTCCCCAGATG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        418 TGTTTGCTCCTTTTACTGGGTGTCTATGGGATGGGGAT-TTTGGGGCTGTGGCTCATATG 476

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||:::|||::::::||||||   ||||||||||||||||||||||||||||||
Db        477 GGAAACATAATGTTTATGTCCTTTTGTGGAGACAACCTTGTCAATCACTATATGTGTGAC 536

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   ||||||||||:::||||||:::|||
Db        537 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATAAATTTGCTGGTGGTTTTT 596

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||:::|||||||||||||||||||||||||||||||   :::|||:::||||||||||||
Db        597 ATTATTGTGACCGTTGGCATTGGGGTGCCGATTGTCACCATTTTTCTCTCTTATGGTTTT 656

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   :::||||||   ||||||||||||||||||||||||:::
Db        657 ATTCTTTCCAGCATTCTCCACATTAGTTCCACAGAGGGCAGGTCTAAAGCCTTCAGTACC 716

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||   ||||||
Db        717 TGCAGTTCCCACATAATTGTGGTATCGCTTTTCTTTGGGTCAGGTGCTTTCATGTACCTC 776

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||   |||
Db        777 AAACCACCTTCTATTCTACCCCTGGACCAGGGGAAAGTGTCCTCCATTTTTTGTACTGCT 836

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        837 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAAGATGTCAAAGTTGCC 896

Qy        301 LeuLysArgThrPheSerArg 307
              |||:::|||||||||   |||
Db        897 CTGAGGAGAACCTTTTGCAGA 917
```

RESULT 5
AAH32421
ID AAH32421 standard; DNA; 930 BP.
XX
AC AAH32421;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 994.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027582.
XX
PR 08-OCT-1999; 99US-0158615P.
PR 24-FEB-2000; 2000US-0184809P.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists.
XX
PS Claim 8; Page 577; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides which
CC encode polypeptides involved in olfactory sensation. The polynucleotides

```
CC   can be used in screening for olfactory agonists and antagonists. The
CC   methods allow for the determination of primary scents and the
CC   identification of the odour receptors used to detect these primary
CC   scents. The methods also enable determination of secondary scents and the
CC   identification of combinations of odour receptors that are involved in
CC   detecting such secondary scents. This enables the construction of a scent
CC   representation (also called a scent fingerprint or scent profile), which
CC   may be used to re-create and edit scents. Libraries of olfactory
CC   receptors are useful for determining the interaction pattern of a
CC   composition with the receptors, and can be used for determining
CC   differences in the olfactory faculties of different individuals
XX
SQ   Sequence 930 BP; 190 A; 251 C; 188 G; 301 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.07e-109       Length:         930
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     4               Gaps:           1

US-10-023-597-24 (1-311) x AAH32421 (1-930)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||::::::|||:::|||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::|||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||   ||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   |||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||   ||||||||||||      ||||||:::||||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   |||||||||||||||||||||||||||||||:::||||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   ||||||   ||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   |||||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   |||||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||   |||   ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 6
AAS15910
ID   AAS15910 standard; cDNA; 933 BP.
XX
AC   AAS15910;
XX
DT   25-JAN-2002  (first entry)
XX
DE   DNA encoding G-protein coupled receptor (GCREC) #14.
XX
KW   G-protein coupled receptor; GCREC; vaccine; gene therapy;
KW   cell proliferation disorder; cancer; arteriosclerosis;
KW   neurological disorder; epilepsy; stroke; cardiovascular disorder;
KW   hypertension; ischaemic heart disease; gastrointestinal disorder;
KW   anorexia; peptic ulcer; autoimmune disorder; inflammatory disorder;
KW   diabetes; osteoporosis; psoriasis; metabolic disorder; obesity;
KW   schizophrenic disorder; neuroskeletal disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .933
FT                   /*tag=  a
FT                   /product= "GCREC 14"
FT                   /note= "G-protein coupled receptor 14"
XX
PN   WO200166742-A2.
XX
PD   13-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US006814.
XX
PR   03-MAR-2000; 2000US-0186854P.
PR   10-MAR-2000; 2000US-0188384P.
PR   17-MAR-2000; 2000US-0190453P.
PR   20-MAR-2000; 2000US-0190730P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lal P,  Tang YT,  Patterson C,  Yao MG,  Shih LL,  Tribouley CM;
PI   Lu DAM,  Yue H,  Khan FA,  Policky JL,  Au-Young J,  Yang J,  Harland L;
PI   Walsh RT,  Lo TP,  Borowsky ML;
XX
DR   WPI; 2001-656776/75.
DR   P-PSDB; AAU10313.
XX
PT   Novel G-protein coupled receptor polypeptides, for treating and
PT   preventing autoimmune/inflammatory disorders, neurological disorders,
PT   cell proliferative disorders, cardiovascular disorders and viral
PT   infections.
XX
PS   Claim 5; Page 138; 141pp; English.
XX
CC   The invention describes a novel isolated polypeptide, selected from a
CC   group of 21 G-protein coupled receptor polypeptides (GCREC) and useful in
CC   vaccines and gene therapy. The polypeptide (I) is useful for screening
CC   for agonist or antagonist of (I), compounds specifically binding to (I),
CC   or compounds that modulate the activity of (I). The polynucleotide
```

```
CC   encoding (1) is useful for screening a compound for effectiveness in
CC   altering expression of a target polynucleotide comprising (II), by
CC   exposing a sample comprising the target polynucleotide to a compound,
CC   detecting altered expression of the target polynucleotide, and comparing
CC   the expression of the target polynucleotide in the presence of varying
CC   amounts of compound and in the absence of the compound. (I) and (II) are
CC   useful for diagnosis, treatment and prevention of cell proliferative
CC   disorders (e.g. cancers, arteriosclerosis, atherosclerosis), neurological
CC   disorders (e.g. epilepsy, stroke, schizophrenic disorders and
CC   neuroskeletal disorders), cardiovascular disorders (e.g. hypertension,
CC   ischaemic heart disease), gastrointestinal disorders (e.g. anorexia,
CC   peptic ulcer), autoimmune/inflammatory disorders (e.g. diabetes mellitus,
CC   osteoporosis, psoriasis), and metabolic disorders such as obesity.
CC   Furthermore, the polynucleotide is useful; as primers for detecting
CC   single nucleotide polymorphisms; as elements in microarray, to monitor or
CC   measure protein-protein interactions, drug-target interactions, and gene
CC   expression profiles; to generate a transcript image of a tissue or cell
CC   type, and to generate hybridisation probes useful in mapping the
CC   naturally occurring genomic sequence. This sequence encodes G-protein
CC   coupled receptor 14, one of 21 GCREC proteins described in the method of
CC   the invention
XX
SQ   Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.09e-109      Length:          933
Score:                    1273.50        Matches:         250
Percent Similarity:       87.46%         Conservative:    22
Best Local Similarity:    80.39%         Mismatches:      38
Query Match:              80.75%         Indels:          1
DB:                       5              Gaps:            1

US-10-023-597-24 (1-311) x AAS15910 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||::::::|||:::|||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||||:::|||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              |||||||||||||||||||   |||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||   ||||||||||||      ||||||:::|||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||||:::|||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   ||||||   ||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::|||||   |||   ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 7
AAS42233
ID   AAS42233 standard; cDNA; 933 BP.
XX
AC   AAS42233;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human cDNA encoding olfactory receptor AOLFR26.
XX
KW   Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW   ss; food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS   Homo sapiens.
XX
PN   WO200168805-A2.
XX
PD   20-SEP-2001.
XX
PF   13-MAR-2001; 2001WO-US007771.
XX
PR   13-MAR-2000; 2000US-0188914P.
PR   24-MAR-2000; 2000US-0192033P.
PR   12-APR-2000; 2000US-0198474P.
PR   24-APR-2000; 2000US-0199335P.
PR   26-MAY-2000; 2000US-0207702P.
PR   23-JUN-2000; 2000US-0213849P.
PR   16-AUG-2000; 2000US-0226534P.
PR   07-SEP-2000; 2000US-0230732P.
PR   07-FEB-2001; 2001US-0266862P.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Zozulya S;
XX
DR   WPI; 2001-570867/64.
DR   P-PSDB; AAU24540.
XX
PT   Nucleic acids encoding human olfactory G protein-coupled receptors,
PT   useful for screening for compounds involved in olfactory sensation, where
PT   the compounds can be used in the food, pharmaceutical and cosmetic
PT   industries to customize odors.
XX
PS   Claim 1; Page 94; 319pp; English.
XX
CC   The invention relates to nucleic acids encoding human olfactory
CC   receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC   specifically recognise molecules, odourants, that elicit specific
```

```
CC   olfactory sensation. The human olfactory receptors and polynucleotides
CC   encoding them are useful for screening a library of chemical compounds
CC   for compounds that are involved in olfactory sensation. Modulators of
CC   their activity are useful for pharmacological and genetic modulation of
CC   olfactory signalling pathways. Therefore, they can be used in the food,
CC   pharmaceutical and cosmetic industries to customise odours and
CC   fragrances. The present sequence encodes a human olfactory receptor of
CC   the invention
XX
SQ   Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.09e-109      Length:          933
Score:                    1273.50        Matches:         250
Percent Similarity:       87.46%         Conservative:    22
Best Local Similarity:    80.39%         Mismatches:      38
Query Match:              80.75%         Indels:          1
DB:                       5              Gaps:            1

US-10-023-597-24 (1-311) x AAS42233 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||::::::|||:::|||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||   ||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||   ||||||||||||   ||||||:::||||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||:::||||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::||||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||      ||||||   ||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::|||||||:::   ||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||   |||   ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 8
ABZ43067
ID   ABZ43067 standard; DNA; 933 BP.
XX
AC   ABZ43067;
XX
DT   06-MAR-2003  (first entry)
XX
DE   Human GPCR polynucleotide SEQ ID NO 395.
XX
KW   Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
KW   drug development; gustatory; taste; fragrance; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200216548-A2.
XX
PD   28-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-IB001446.
XX
PR   04-AUG-2000; 2000JP-00237818.
PR   13-FEB-2001; 2001JP-00034434.
XX
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI   Haga T,  Takeda S,  Mitaku S;
XX
DR   WPI; 2002-304118/34.
DR   P-PSDB; ABP95793.
XX
PT   Database global search for G protein-coupled receptors, proteins and
PT   encoded genes for studying in vivo signal transduction mechanism and
PT   identifying targets for drug development.
XX
PS   Claim 9; SEQ ID NO 395; 97pp + Sequence Listing; Japanese.
XX
CC   The invention relates to a method for screening G protein-coupled
CC   receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
CC   ABP95942) by extracting open-reading frames containing 6-8 transmembrane
CC   domains with 250-1000 amino acid residues to give a gene homologous with
CC   a known GPCR gene. The receptor proteins and encoded genes are useful for
CC   studying in vivo signal transduction mechanism and identifying targets
CC   for drug development e.g. based on olfactory and gustatory receptors in
CC   form of agonists and antagonists by screening intrinsic and extrinsic
CC   ligands as bitter taste inhibitors, taste enhancers and fragrance
CC   improvers. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.09e-109      Length:          933
Score:                    1273.50        Matches:         250
Percent Similarity:       87.46%         Conservative:    22
Best Local Similarity:    80.39%         Mismatches:      38
Query Match:              80.75%         Indels:          1
DB:                       6              Gaps:            1

US-10-023-597-24 (1-311) x ABZ43067 (1-933)
```

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   ||||||||||||||||||||||||    ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::::|||:::||||||||||||||||||||||    ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 ||||||||||||||||||||   ||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||   ||||||||||||      ||||||:::|||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||:::|||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||      ||||||   ||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::|||||||:::   ||||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||      |||      ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930
```

RESULT 9
ABK68455
ID ABK68455 standard; DNA; 933 BP.
XX
AC ABK68455;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human DNA for olfactory and pheromone G protein-coupled receptor #55.
XX
KW Human; ds; gene; olfactory and pheromone G protein coupled receptor;
KW GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;
KW anorectic; taste; fragrance; food additive; cosmetic; cell migration;
KW sterility; psychotic disorder; neurological disorder; anxiety;
KW schizophrenia; manic depression; depression; axonal growth;
KW menstrual cycle; appetite sexual motivation; sexual attraction;
KW aggression.
XX
OS Homo sapiens.
XX
PN WO200224726-A2.
XX
PD 28-MAR-2002.
XX
PF 21-SEP-2001; 2001WO-BE000162.
XX
PR 22-SEP-2000; 2000EP-00870211.
XX
PA (CHEM-) CHEMCOM SA.
XX
PI Veithen A;
XX
DR WPI; 2002-330013/36.
DR P-PSDB; AAU95568.
XX
PT Novel pheromone G-protein coupled receptor and receptor-derived agonists,
PT antagonists or inhibitors useful in food or cosmetic products or in the
PT treatment or prevention of neurological disorders such as anxiety and
PT schizophrenia.
XX
PS Disclosure; Page 178-179; 833pp; English.
XX
CC The invention relates to olfactory and Pheromone G-protein coupled
CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active
CC portion and its encoding polynucleotide. Also included are an agonist,
CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector
CC comprising the polynucleotide, a cell transformed by the vector, a non-
CC human mammal comprising a partial or total deletion of the polynucleotide
CC encoding the receptor and screening (detection and possibly, recovering)
CC of compounds which are known or not known to be agonist, antagonists or
CC inhibitors of natural compounds to the GPCR. The receptor-derived
CC agonists, antagonists, inhibitors or compounds are used as an
CC improvement, elimination or substitution of an existing taste and/or a
CC fragrance of (or in) the food and/or cosmetic products. They can also be
CC used in the preparation of medicament in the treatment and/or prevention
CC of a mammalian disorder, such as cell migration, sterility, psychotic and
CC neurological disorders, including anxiety, schizophrenia, manic
CC depression, depression, for promoting axonal growth, nerve cell
CC connection and nerve regeneration for modulating male and female
CC endocrine functions, hormone production and the menstrual cycle, for the
CC prevention or the treatment by stimulation of several mammalian
CC behaviours, such as stimulation or suppression of appetite, sexual
CC motivation, sexual attraction, aggression and for promoting or
CC suppressing chemical communication between organisms. The present
CC sequence is a human DNA encoding an olfactory and pheromone GPCR
XX
SQ Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.09e-109	Length:	933
Score:	1273.50	Matches:	250
Percent Similarity:	87.46%	Conservative:	22
Best Local Similarity:	80.39%	Mismatches:	38
Query Match:	80.75%	Indels:	1
DB:	6	Gaps:	1

US-10-023-597-24 (1-311) x ABK68455 (1-933)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   ||||||||||||||||||||||||    ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57
```

```
Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::|||:::|||||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||   ||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||   ||||||||||||       ||||||:::|||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   ||||||||||||||||||||||||||||||:::||||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||   ||||||   ||||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||      |||      ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 10
ABK37521
ID   ABK37521 standard; cDNA; 933 BP.
XX
AC   ABK37521;
XX
DT   08-MAY-2002  (first entry)
XX
DE   DNA encoding G-coupled olfactory receptor #23.
XX
KW   Human; olfactory G-coupled receptor; sensory perception of odourant;
KW   odour composition; taste composition; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200198526-A2.
XX
PD   27-DEC-2001.
XX
PF   22-JUN-2001; 2001WO-US020122.
XX
PR   22-JUN-2000; 2000US-0213812P.
PR   13-MAR-2001; 2001US-00804291.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Zozulya S,  Stryer L;
XX
DR   WPI; 2002-083330/11.
DR   P-PSDB; AAU85162.
XX
PT   Representing sensory perception of one or more odorants for the
PT   identification and design of tastes and odors comprises providing a
PT   representative group of n olfactory receptors.
XX
PS   Example; Page 68; 182pp; English.
XX
CC   The invention relates to a method of representing sensory perception of
CC   one or more odourants. The method comprises: (a) providing a
CC   representative class of n olfactory receptors or ligand binding domains
CC   (LBD) of these receptors; (b) measuring values X1 to Xn representative of
CC   at least one activity of one or more odourants selected from: (i) binding
CC   one or more odourants to the LBD of at least one of the n olfactory
CC   receptors; (ii) activating at least one of the n olfactory receptors with
CC   the one or more odourants; and (iii) blocking at least one of the n
CC   olfactory receptors with the one or more odourants; and (c) generating a
CC   representation of sensory perception from the values X1 to Xn. The
CC   representation of the sensory perception of odourants is useful for the
CC   design and formulation of odour and taste compositions. ABK37499-ABK37754
CC   and ABK37918-ABK37921 represent human olfactory G-coupled receptor coding
CC   sequences and related PCR primers of the invention
XX
SQ   Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.09e-109       Length:         933
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     6               Gaps:           1

US-10-023-597-24 (1-311) x ABK37521 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   ||||||||||||||||||||||   |||||||   |||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::|||:::|||||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
```

```
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930
```

RESULT 11
ABS58782
ID ABS58782 standard; DNA; 991 BP.
XX
AC ABS58782;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human G-protein coupled receptor (GPCR) gene #36.
XX
KW Human; G-protein coupled receptor; GPCR; cardiomyopathy; atherosclerosis;
KW diabetes; cell signal processing; metabolic pathway modulation; cancer;
KW adenocarcinoma; lymphoma; prostate cancer; uterus cancer; asthma;
KW immune response; neurodegenerative disorder; inflammatory disorder;
KW Crohn's disease; multiple sclerosis; Albright hereditary osteodystrophy;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200259313-A2.
XX
PD 01-AUG-2002.
XX
PF 18-DEC-2001; 2001WO-US049394.
XX
PR 18-DEC-2000; 2000US-0256635P.
PR 21-DEC-2000; 2000US-0257876P.
PR 04-JAN-2001; 2001US-0259743P.
PR 10-JAN-2001; 2001US-0260718P.
PR 12-JAN-2001; 2001US-0261498P.
PR 24-JAN-2001; 2001US-0263689P.
PR 08-FEB-2001; 2001US-0267464P.
PR 22-FEB-2001; 2001US-0271021P.
PR 14-MAR-2001; 2001US-0275946P.
PR 23-MAR-2001; 2001US-0278150P.
PR 18-APR-2001; 2001US-0284591P.
PR 23-APR-2001; 2001US-0285718P.
PR 19-JUN-2001; 2001US-0299327P.
PR 16-AUG-2001; 2001US-0312902P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Li L, Ballinger RA, Padigaru M, Kekuda R, Colman SD, Spytek KA;
PI Casman SJ, Vernet CAM, Shenoy SG, Gusev V, Malyankar UM, Edinger S;
PI Gerlach V, Smithson G, Stone DJ, Sciore P, Macdougall JR, Gunther E;
PI Peyman JA, Ellerman K, Gangolli EA, Millet I;
XX
DR WPI; 2002-599789/64.
DR P-PSDB; ABG76802.
XX
PT New G protein coupled receptor polypeptides and polynucleotides, useful
PT in gene therapy, particularly for treating or preventing cardiomyopathy,
PT atherosclerosis, diabetes, multiple sclerosis, Crohn's disease or cancer
PT in humans.
XX
PS Claim 1; Page 118; 685pp; English.
XX
CC The invention relates to novel isolated G-protein coupled receptor (GPCR)
CC polypeptides and polynucleotides. The GPCR polypeptide, GPCR nucleic acid
CC and antibody are useful for treating, preventing or alleviating a GPCR-
CC associated disorder or a pathological state in a subject, particularly a
CC human. In particular, the disorder is cardiomyopathy, atherosclerosis,
CC diabetes, or a disorder related to cell signal processing and metabolic
CC pathway modulation. The GPCR polypeptide and nucleic acid are also useful
CC for diagnosing the presence of or predisposition to a disease associated
CC with altered levels of GPCR, particularly cancer. The GPCR nucleic acid
CC and polypeptide are especially useful in therapeutic or prophylactic
CC applications for disorders associated with aberrant GPCR expression or
CC activity. The DNA encoding the protein is useful in gene therapy for
CC treating the above conditions. Furthermore, the nucleic acids and
CC polypeptides are useful in treating adenocarcinoma, lymphoma, prostate
CC cancer, uterus cancer, immune response, neurodegenerative disorders,
CC asthma, inflammatory disorders, Crohn's disease, multiple sclerosis or
CC Albright hereditary osteodystrophy. These are also useful in developing a
CC powerful assay system for functional analysis of various human disorders,
CC as well as in diagnostic applications. ABS58747-ABS59231 represent human
CC GPCR coding sequences, primers and probes of the invention
XX
SQ Sequence 991 BP; 205 A; 270 C; 198 G; 318 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.57e-109	Length:	991
Score:	1273.50	Matches:	250
Percent Similarity:	87.46%	Conservative:	22
Best Local Similarity:	80.39%	Mismatches:	38
Query Match:	80.75%	Indels:	1
DB:	6	Gaps:	1

US-10-023-597-24 (1-311) x ABS58782 (1-991)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db         49 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 105

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db        106 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 165

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        166 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 225

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
```

```
Db        226 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 285

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        286 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 345

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        346 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 405

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        406 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 465

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        466 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 525

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        526 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 585

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        586 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 645

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        646 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 705

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        706 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 765

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        766 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 825

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        826 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 885

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        886 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 945

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db        946 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 978

RESULT 12
ADC85838
ID   ADC85838 standard; DNA; 1333 BP.
XX
AC   ADC85838;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human GPCR gene SEQ ID NO:291.
XX
KW   ds; gene; human; GPCR;
KW   guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1270724-A2.
XX
PD   02-JAN-2003.
XX
PF   18-JUN-2002; 2002EP-00013517.
XX
PR   18-JUN-2001; 2001JP-00246789.
XX
PA   (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Suwa M,  Asai K,  Akiyama Y,  Aburatani H;
XX
DR   WPI; 2003-315783/31.
DR   P-PSDB; ADC85839.
XX
PT   New polynucleotide, useful for preparing a composition for treating a
PT   patient in need of increased or suppressed activity or expression of the
PT   guanosine triphosphate-binding protein coupled receptor.
XX
PS   Claim 1; SEQ ID NO 291; 28pp; English.
XX
CC   The invention relates to a novel polynucleotide encoding a guanosine
CC   triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC   the invention may have a use in gene therapy. The polynucleotide and
CC   polypeptide are useful for preparing a composition for treating a patient
CC   in need of increased or suppressed activity or expression of the
CC   guanosine triphosphate-binding protein coupled receptor. The
CC   polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC   invention.
XX
SQ   Sequence 1333 BP; 296 A; 327 C; 268 G; 442 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                  9.5e-109        Length:          1333
Score:                      1273.50         Matches:         250
Percent Similarity:         87.46%          Conservative:    22
Best Local Similarity:      80.39%          Mismatches:      38
Query Match:                80.75%          Indels:          1
DB:                         9               Gaps:            1

US-10-023-597-24 (1-311) x ADC85838 (1-1333)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db        201 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 257

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db        258 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 317

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        318 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 377

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db        378 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 437

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        438 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 497

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        498 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 557

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        558 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 617

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        618 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 677

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        678 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 737

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
```

```
Db        738 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 797

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        798 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 857

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        858 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 917

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        918 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 977

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        978 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 1037

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db       1038 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 1097

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db       1098 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 1130
```

RESULT 13
ABK97207
ID ABK97207 standard; cDNA; 976 BP.
XX
AC ABK97207;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human G-protein coupled receptor (GPCR) GPCRX gene #1.
XX
KW G-protein coupled receptor; receptor; GPCR; GPCRX; cardiomyopathy;
KW atherosclerosis; diabetes; cell signal processing; cancer; trauma;
KW metabolic pathway modulation; neuro-olfactory system; surgery;
KW neoplastic disorder; adenocarcinoma; lymphoma; prostate cancer;
KW uterus cancer; immune response; acquired immunodeficiency syndrome; AIDS;
KW asthma; Crohn's disease; multiple sclerosis;
KW Albright hereditary osteodystrophy; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200250117-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US049077.
XX
PR 18-DEC-2000; 2000US-0256635P.
PR 21-DEC-2000; 2000US-0257876P.
PR 04-JAN-2001; 2001US-0259743P.
PR 10-JAN-2001; 2001US-0260718P.
PR 12-JAN-2001; 2001US-0261498P.
PR 24-JAN-2001; 2001US-0263689P.
PR 08-FEB-2001; 2001US-0267464P.
PR 22-FEB-2001; 2001US-0271021P.
PR 14-MAR-2001; 2001US-0275946P.
PR 23-MAR-2001; 2001US-0278150P.
PR 18-APR-2001; 2001US-0284591P.
PR 23-APR-2001; 2001US-0285718P.
PR 19-JUN-2001; 2001US-0299327P.
PR 16-AUG-2001; 2001US-0312902P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;
PI Vernet CAM, Li L, Shenoy S, Casman SJ;
XX
DR WPI; 2002-528447/56.
DR P-PSDB; ABG68134.
XX
PT New G-protein coupled receptor polypeptides for treating or preventing
PT cardiomyopathy, atherosclerosis, diabetes, multiple sclerosis, acquired
PT immunodeficiency syndrome or cancer in humans.
XX
PS Claim 5; Page 79; 110pp; English.
XX
CC The present invention relates to a new G-protein coupled receptor (GPCRX)
CC polypeptide. The GPCRX polypeptide, GPCRX nucleic acid and antibody are
CC useful for treating, preventing or alleviating a GPCRX-associated
CC disorder or a pathological state in a subject, particularly a human. In
CC particular, the disorder is cardiomyopathy, atherosclerosis, diabetes, or
CC a disorder related to cell signal processing and metabolic pathway
CC modulation. The GPCRX polypeptide and nucleic acid are also useful for
CC diagnosing the presence of or predisposition to a disease associated with
CC altered levels of GPCRX, particularly cancer. The GPCRX nucleic acid and
CC polypeptide are especially useful in therapeutic or prophylactic
CC applications for disorders of the neuro-olfactory system, e.g. those
CC induced by trauma, surgery and/or neoplastic disorders. The DNA encoding
CC the protein is useful in gene therapy for treating the above conditions.
CC Furthermore, the nucleic acids and polypeptides are useful in treating
CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune
CC response, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's
CC disease, multiple sclerosis or Albright hereditary osteodystrophy. These
CC are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications. The present nucleic acid sequence represents one of a
CC collection (ABK97207-ABK97226) of human GPCRX genes that encode the human
CC GPCRX proteins (AAU68134-AAU68153) of the invention
XX
SQ Sequence 976 BP; 204 A; 259 C; 204 G; 309 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.09e-108	Length:	976
Score:	1268.00	Matches:	247
Percent Similarity:	86.50%	Conservative:	22
Best Local Similarity:	79.42%	Mismatches:	42
Query Match:	80.41%	Indels:	0
DB:	6	Gaps:	0

US-10-023-597-24 (1-311) x ABK97207 (1-976)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db          8 ATGGGCCAGCCCAAAAACTCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 67

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db         68 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 127

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        128 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 187

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db        188 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 247

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        248 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 307

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        308 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 367

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        368 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 427

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
```

```
Db        428 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 487

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        488 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 547

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        548 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 607

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        608 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 667

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        668 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 727

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        728 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 787

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        788 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 847

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        848 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 907

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db        908 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 940

RESULT 14
AAD60328
ID   AAD60328 standard; DNA; 976 BP.
XX
AC   AAD60328;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human G-protein coupled receptor (GPCR) GPCRX DNA #1.
XX
KW   Human; G-protein coupled receptor; GPCRX; stroke; obesity; virucide;
KW   multiple sclerosis; Alzheimer's disease; graft-versus-host disease;
KW   endometriosis; tissue typing; gene therapy; vaccine; antibacterial;
KW   retinal disease; infectious disease; bulimia; receptor; fungicide;
KW   Parkinson's disease; hypertension; acute heart failure; infection;
KW   cardiomyopathy; atherosclerosis; diabetes; detectability disorder;
KW   cancer; signal transduction pathway disorder; metabolic disorder;
KW   developmental disorder; osteoporosis; protozoacide; asthma; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             8. .943
FT                   /*tag=  a
FT                   /product= "Human GPCRX protein"
XX
PN   US2003100491-A1.
XX
PD   29-MAY-2003.
XX
PF   18-DEC-2001; 2001US-00024399.
XX
PR   18-DEC-2000; 2000US-0256635P.
PR   21-DEC-2000; 2000US-0257876P.
PR   04-JAN-2001; 2001US-0259743P.
PR   10-JAN-2001; 2001US-0260718P.
PR   12-JAN-2001; 2001US-0261498P.
PR   24-JAN-2001; 2001US-0263689P.
PR   08-FEB-2001; 2001US-0267464P.
PR   22-FEB-2001; 2001US-0271021P.
PR   14-MAR-2001; 2001US-0275946P.
PR   23-MAR-2001; 2001US-0278150P.
PR   18-APR-2001; 2001US-0284591P.
PR   23-APR-2001; 2001US-0285718P.
PR   19-JUN-2001; 2001US-0299327P.
PR   16-AUG-2001; 2001US-0312902P.
XX
PA   (PADI/) PADIGARU M.
PA   (KEKU/) KEKUDA R.
PA   (COLM/) COLMAN S D.
PA   (SPYT/) SPYTEK K A.
PA   (BALL/) BALLINGER R A.
PA   (VERN/) VERNET C A M.
PA   (LILL/) LI L.
PA   (SHEN/) SHENOY S G.
PA   (CASM/) CASMAN S J.
XX
PI   Padigaru M,  Kekuda R,  Colman SD,  Spytek KA,  Ballinger RA;
PI   Vernet CAM,  Li L,  Shenoy SG,  Casman SJ;
XX
DR   WPI; 2003-755210/71.
DR   P-PSDB; AAE39659.
XX
PT   New G-protein coupled receptor (GPCR) polypeptides and nucleic acids,
PT   useful for diagnosing, preventing or treating GPCR-associated disorders,
PT   e.g. cardiomyopathy, atherosclerosis, cancer or diabetes, and in
PT   pharmacogenomics.
XX
PS   Claim 8; Page 32; 46pp; English.
XX
CC   The invention relates to G-protein coupled receptor (GPCR) polypeptides
CC   (designated as GPCRX) and nucleic acid sequences. GPCRX polypeptides,
CC   nucleic acids and antibodies are useful in diagnosing, preventing or
CC   treating GPCRX-associated disorders, such as cardiomyopathy, diabetes,
CC   atherosclerosis and cancer. These may also be used for treating or
CC   preventing other diseases like developmental disorders, taste and scent
CC   detectability disorders, signal transduction pathway disorders, retinal
CC   diseases, metabolic disorders (e.g. obesity), infectious diseases (e.g.
CC   bacterial, fungal, protozoal or viral infections), bulimia, asthma,
CC   Parkinson's disease, hypertension, acute heart failure, osteoporosis,
CC   multiple sclerosis, Alzheimer's disease, stroke, graft-versus-host
CC   disease or endometriosis. GPCRX nucleic acid is used as hybridisation
CC   probes, in chromosome mapping, tissue typing, preventive medicine and
CC   pharmacogenomics. It is also useful in gene therapy. GPCRX is useful as
CC   vaccines. The present sequence is human GPCRX DNA
XX
SQ   Sequence 976 BP; 204 A; 259 C; 204 G; 309 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.09e-108      Length:          976
Score:                    1268.00        Matches:         247
Percent Similarity:       86.50%         Conservative:    22
Best Local Similarity:    79.42%         Mismatches:      42
Query Match:              80.41%         Indels:          0
DB:                       9              Gaps:            0

US-10-023-597-24 (1-311) x AAD60328 (1-976)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db          8 ATGGGCCAGCCCAAAAACTCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 67

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db         68 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 127

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        128 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 187

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
```

```
Db        188 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 247

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        248 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 307

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        308 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 367

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        368 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 427

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        428 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 487

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        488 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 547

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        548 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 607

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        608 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 667

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        668 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 727

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        728 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 787

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        788 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 847

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        848 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 907

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db        908 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 940
```

RESULT 15
ABT05655
ID ABT05655 standard; DNA; 953 BP.
XX
AC ABT05655;
XX
DT 16-OCT-2002 (first entry)
XX
DE GPCR 4 protein encoding DNA SEQ ID No 13.
XX
KW Antidiabetic; cytostatic; anorectic; nootropic; neuroprotective; GPCRX;
KW antiparkinsonian; cardiant; antiarteriosclerotic; immunosuppressive;
KW hypotensive; haemostatic; antifertility; antiasthmatic; antiinflammatory;
KW anti-HIV; G-protein coupled receptor X; cardiomyopathy; atherosclerosis;
KW diabetes; metabolic disorder; diabetes; obesity; infectious disease;
KW anorexia; neurodegenerative disorder; Alzheimer's disease; dyslipidaemia;
KW Parkinson's disorder; immune disorder; haematopoietic disorder; obesity;
KW metabolic syndrome X; wasting disorder; cancer; hypertension; neoplasm;
KW congenital heart defect; aortic stenosis; subaortic stenosis; lymphoma;
KW transplantation; adrenoleukodystrophy; congenital adrenal hyperplasia;
KW prostate cancer; adenocarcinoma; uterus cancer; fertility; haemophilia;
KW hypercoagulation; idiopathic thrombocytopenic purpura; immunodeficiency;
KW graft versus host disease; AIDS; bronchial asthma; Crohn's disease;
KW multiple sclerosis; chromosomal mapping; tissue typing; forensic biology;
KW gene therapy; transgenic animal; gene; ds.
XX
OS Unidentified.
XX
PN WO200246229-A2.
XX
PD 13-JUN-2002.
XX
PF 05-DEC-2001; 2001WO-US046530.
XX
PR 05-DEC-2000; 2000US-0251459P.
PR 29-DEC-2000; 2000US-0259007P.
PR 04-DEC-2001; 2001US-00005041.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Casman SJ, Padigaru M, Burgess CE, Shimkets RA, Spytek KA;
PI Gilbert JA, Mayotte JE, Baumgartner JC, Mishra V, Vernet CAM;
PI Dickinson KS, Ballinger RA, Wolenc AR;
XX
DR WPI; 2002-537559/57.
DR P-PSDB; ABJ04724.
XX
PT Novel isolated G-protein coupled receptor polypeptide, designated NOVX,
PT useful for treating or preventing in human receptor-associated disorders
PT e.g. cardiomyopathy, atherosclerosis or diabetes.
XX
PS Claim 9; Page 33; 264pp; English.
XX
CC The invention relates to a novel isolated G-protein coupled receptor X
CC (GPCRX) polypeptide. The isolated protein, its encoding polynucleotide
CC and the antibody of the isolated protein is useful for treating or
CC preventing a GPCRX-associated disorder in a subject, preferably human,
CC where the disorder is cardiomyopathy, atherosclerosis or diabetes. The
CC isolated protein, its encoding polynucleotide and the antibody of the
CC isolated protein is useful for treating or preventing metabolic
CC disorders, diabetes, obesity, infectious disease, anorexia,
CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC immune disorders, haematopoietic disorders, and various dyslipidaemias,
CC metabolic disturbances associated with obesity, the metabolic syndrome X,
CC wasting disorders associated with chronic diseases, and cancer. The
CC disorders also include cardiomyopathy, atherosclerosis, hypertension,
CC congenital heart defects, aortic stenosis, subaortic stenosis,
CC transplantation, adrenoleukodystrophy, congenital adrenal hyperplasia,
CC prostate cancer, neoplasm, adenocarcinoma, lymphoma, uterus cancer,
CC fertility, haemophilia, hypercoagulation, idiopathic thrombocytopenic
CC purpura, immunodeficiencies, graft versus host disease, AIDS, bronchial
CC asthma, Crohn's disease, and multiple sclerosis. The isolated protein,
CC its encoding polynucleotide and the antibody of the isolated protein is
CC useful in screening assays, detection assays (e.g., chromosomal mapping,
CC tissue typing, forensic biology). The isolated polynucleotide is useful
CC in gene therapy, to express the isolated protein, to detect GPCRX mRNA or
CC a genetic lesion in a GPCRX gene, and to modulate GPCRX activity. The
CC cell of the invention is useful for producing non-human transgenic
CC animals. This polynucleotide sequence represents DNA encoding a GPCRX
CC protein of the invention
XX
SQ Sequence 953 BP; 202 A; 256 C; 193 G; 302 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.79e-108	Length:	953
Score:	1266.50	Matches:	249
Percent Similarity:	87.14%	Conservative:	22
Best Local Similarity:	80.06%	Mismatches:	39
Query Match:	80.31%	Indels:	1
DB:	6	Gaps:	1

US-10-023-597-24 (1-311) x ABT05655 (1-953)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
```

```
Db         16 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 72

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db         73 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 132

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        133 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 192

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db        193 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 252

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        253 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 312

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        313 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 372

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        373 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 432

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        433 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 492

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        493 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 552

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        553 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCCGGTGGTCTTT 612

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        613 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 672

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        673 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 732

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        733 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 792

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        793 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 852

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        853 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 912

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db        913 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 945
```

Search completed: September 30, 2004, 10:09:15
Job time : 455 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 09:52:11 ; Search time 86 Seconds
(without alignments)
2006.857 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 1577
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163410_14903/app_query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023597_@CGN_1_1_69_@runat_29092004_163410_14903 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	752	47.7	485	4	US-09-621-976-18883	Sequence 18883, A
2	705.5	44.7	945	4	US-09-016-434-1114	Sequence 1114, Ap
3	630	39.9	1282	4	US-09-016-434-1413	Sequence 1413, Ap
4	621	39.4	1062	4	US-09-668-680-10	Sequence 10, Appl
5	619	39.3	1854	4	US-09-016-434-1312	Sequence 1312, Ap
6	616	39.1	966	3	US-08-748-506-6	Sequence 6, Appli
7	614	38.9	966	3	US-08-748-506-5	Sequence 5, Appli
8	608	38.6	966	3	US-08-748-506-8	Sequence 8, Appli
9	606	38.4	1351	4	US-09-546-986A-5	Sequence 5, Appli
10	606	38.4	1351	4	US-09-524-730-5	Sequence 5, Appli
11	594	37.7	966	3	US-08-748-506-7	Sequence 7, Appli
12	591.5	37.5	3459	4	US-09-016-434-1363	Sequence 1363, Ap
13	589	37.3	1290	2	US-08-827-291A-1	Sequence 1, Appli
14	584.5	37.1	1438	4	US-09-016-434-1313	Sequence 1313, Ap
15	574	36.4	1035	4	US-09-546-986A-1	Sequence 1, Appli
16	574	36.4	1035	4	US-09-524-730-1	Sequence 1, Appli
17	572.5	36.3	1713	2	US-08-467-948A-1	Sequence 1, Appli
18	572.5	36.3	1713	3	US-08-467-947A-1	Sequence 1, Appli
19	567	36.0	1297	4	US-09-668-680-11	Sequence 11, Appl
20	561	35.6	1990	4	US-09-016-434-1056	Sequence 1056, Ap
21	560	35.5	1080	4	US-09-668-680-9	Sequence 9, Appli
22	552	35.0	1411	4	US-09-546-986A-3	Sequence 3, Appli
23	552	35.0	1411	4	US-09-524-730-3	Sequence 3, Appli
24	542.5	34.4	951	4	US-09-465-901-47	Sequence 47, Appl
25	535.5	34.0	952	4	US-09-016-434-1115	Sequence 1115, Ap
26	527	33.4	900	3	US-09-085-371-5	Sequence 5, Appli
27	525	33.3	1065	4	US-09-546-986A-7	Sequence 7, Appli
28	525	33.3	1065	4	US-09-524-730-7	Sequence 7, Appli
29	519	32.9	669	4	US-09-465-901-37	Sequence 37, Appl
30	488	30.9	669	4	US-09-465-901-35	Sequence 35, Appl
31	487	30.9	984	3	US-08-748-506-9	Sequence 9, Appli
32	465	29.5	669	4	US-09-465-901-15	Sequence 15, Appl
33	464	29.4	669	4	US-09-465-901-13	Sequence 13, Appl
34	460.5	29.2	675	4	US-09-465-901-43	Sequence 43, Appl
35	452	28.7	669	4	US-09-465-901-31	Sequence 31, Appl
36	444	28.2	669	4	US-09-465-901-41	Sequence 41, Appl
37	436	27.6	648	4	US-09-016-434-1374	Sequence 1374, Ap
38	426	27.0	678	4	US-09-465-901-45	Sequence 45, Appl
39	424	26.9	669	4	US-09-465-901-39	Sequence 39, Appl
40	417.5	26.5	1539	4	US-09-668-680-13	Sequence 13, Appl
41	408	25.9	1828	3	US-08-988-876-2	Sequence 2, Appli
42	405	25.7	648	4	US-09-016-434-1376	Sequence 1376, Ap
43	401	25.4	675	4	US-09-465-901-17	Sequence 17, Appl
44	398.5	25.3	675	4	US-09-016-434-1373	Sequence 1373, Ap
45	397	25.2	669	4	US-09-465-901-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-621-976-18883/c
; Sequence 18883, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18883
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-18883

Alignment Scores:

Pred. No.:	1.04e-72	Length:	485
Score:	752.00	Matches:	156
Percent Similarity:	96.32%	Conservative:	1
Best Local Similarity:	95.71%	Mismatches:	5
Query Match:	47.69%	Indels:	3
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x US-09-621-976-18883 (1-485)

```
Qy          61 PhePheProPheAsnLeuSerLeuValAsp-PheSerPheSerThrThrIleIleProLy 80
               ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db         485 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTTAGTTTCTCTACGACCATCATTCCCAA 426

Qy          80 sMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGl 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        425 AATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCA 366

Qy        100 nPhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGl 120
              ||||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db        365 GTTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCAT-CTGTCGGCGATGGCGTA 307

Qy        120 uAspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVa 140
              ||||||||||||    |||||||||||||||||||||||:::|||||||||||||||||
Db        306 GGACCGCTACGTGGCCATCTGTAACCCATTGTTGTACACGGTCACCATGTCTCCCCAGGT 247

Qy        140 lCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisTh 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATAC 187

Qy        160 rGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAs 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGA 127

Qy        180 pIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePh 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTT 67

Qy        200 eIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPh 220
              |||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db         66 TATTGTTGTGACCGTTGGCATTGGGTG-CCCATTGTTGCCGTTTTTATCTCTTATGGTTT 8

Qy        220 eIleLeu 222
              |||||||
Db          7 TATTCTT 1
```

RESULT 2
US-09-016-434-1114
; Sequence 1114, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 945 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1336040
US-09-016-434-1114

Alignment Scores:

Pred. No.:	2.93e-67	Length:	945
Score:	705.50	Matches:	139
Percent Similarity:	65.59%	Conservative:	65
Best Local Similarity:	44.69%	Mismatches:	100
Query Match:	44.74%	Indels:	7
DB:	4	Gaps:	3

US-10-023-597-24 (1-311) x US-09-016-434-1114 (1-945)

```
Qy          5 AsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGln 24
              |||   :::   |||||||||||||||||||   |||         :::|||   ||||||
Db         19 AACTACACGTTGGTCACTGAGTTTATTCTATTAGGTTTTCCAACTCGCCCTGAACTGCAG 78

Qy         25 ValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGly 44
              :::   :::|||   :::||||||      ||||||:::   ::::::||||||:::|||
Db         79 ATTGTCCTGTTCCTCATGTTTCTGACATTGTATGCTATAATTCTGATAGGGAACATTGGA 138

Qy         45 LeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePhePhePhe 64
              |||::::::||||||   ::::::      |||      ||||||||||||||||
Db        139 TTGATGCTGTTGATCAGGATTGATCCTCACCTTCAAACCCCCATGTATTTTTTCCTTAGC 198

Qy         65 AsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMetSer 84
              |||||||||   ||||||      :::   :::   |||:::||||||||||||::::::
Db        199 AACCTATCATTTGTAGACCTTTGCTATTTCTCAGACATTGTTCCCAAAATGCTGGTCAAT 258

Qy         85 PheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePhe 104
              |||:::|||             ||||||:::   ||||||      ||||||:::||||||
Db        259 TTCCTCTCGGAGAACAAATCTATTTCCTATTATGGGTGTGCCCTGCAGTTTTATTTTTTC 318

Qy        105 CysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyrVal 124
              |||   |||      :::||||||||||||||:::||||||      ||||||||||||
Db        319 TGTACTTTTGCAGATACAGAATCCTTCATCCTGGCCGCCATGGCCTATGATCGCTATGTC 378

Qy        125 GlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeu 144
              ||||||||||||||||||||||:::   ||||||      :::|||:::   |||
Db        379 GCCATCTGTAACCCTTTATTGTACACAGTTGTGATGTCTAGGGGCATCTGTATGCGGTTG 438

Qy        145 LeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleVal 164
              ::::::      |||      |||      ::::::   ||||||
Db        439 ATTGTCTTGTCATACCTTGGAGGCAACATGAGTTCCCTGGTTCACACATCCTTTGCCTTT 498

Qy        165 PheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeu 184
              |||   :::|||      |||::::::||||||:::   ||||||:::   ||||||
Db        499 ATTCTGAAATATTGTGACAAAAATGTTATTAATCATTTTTTCTGTGACCTCCCTCCCCTG 558

Qy        185 LeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePheIleValValThr 204
              |||:::|||||||||      :::   ||||||            :::::::::   |||
Db        559 CTTAAACTATCCTGCACTGACACAACAATTAATGAG---------TGGCTCCTCTCCACA 609

Qy        205 ValGlyIleGlyValProIleValAla---------ValPheIleSerTyrGlyPheIle 221
              |||      |||   |||:::             :::   |||||||||   ||||||
Db        610 TACGGCAGCTCAGTGGAAATCATTTGTTTTATCATCATCATCATCTCCTACTTTTTCATT 669

Qy        222 LeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSerCys 241
              |||   |||:::|||::::::   |||      ||||||   |||   ||||||:::|||
Db        670 CTTCTCTCAGTCTTAAAGATCCGCTCTTTCAGTGGGAGGAAGAAGACCTTTTCTACATGC 729

Qy        242 SerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeuLys 261
              :::|||::::::   :::|||:::::::::   |||:::      |||   |||   :::
Db        730 GCCTCTCACCTGACTTCAGTGACGATCTACCAAGGGACTCTCCTCTTTATTTACTCACGG 789

Qy        262 ProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThrVal 281
              |||      :::         :::      |||:::   |||:::||||||||||
Db        790 CCCAGCTACCTGTATTCTCCAAACACTGATAAAATTATCTCAGTGTTCTACACCATTTTC 849

Qy        282 ValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeu 301
              :::|||:::   ||||||||||||||||||||||||||||||||||||||||   |||
```

```
Db        850 ATTCCAGTGCTGAATCCGTTGATTTATAGTTTGAGAAATAAAGATGTAAAGGATGCAGCT 909

Qy        302 LysArgThrPhe---SerArgIleSerPheSer 311
              ::::::      |||::::::      |||
Db        910 GAGAAAGTTCTAAGATCAAAGGTAGATTCTTCA 942

RESULT 3
US-09-016-434-1413
; Sequence 1413, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1413:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1282 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g516319
US-09-016-434-1413

Alignment Scores:
Pred. No.:              6.85e-59        Length:         1282
Score:                  630.00          Matches:        122
Percent Similarity:     60.87%          Conservative:   60
Best Local Similarity:  40.80%          Mismatches:     117
Query Match:            39.95%          Indels:         0
DB:                     4               Gaps:           0

US-10-023-597-24 (1-311) x US-09-016-434-1413 (1-1282)

Qy          5 AsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGln 24
              |||   :::   :::   ||||||||||||   ||||||:::      ||||||||||||
Db        180 AATGGAACAGTCATTGCTGAGTTCATCCTGCTGGGCTTGCTGGAGGCGCCAGGGCTGCAG 239

Qy         25 ValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGly 44
                ||||||   |||||||||      |||   |||||||||   |||||||||
Db        240 CCAGTTGTCTTTGTGCTCTTCCTCTTTGCCTACCTGGTCACGGTCAGGGGCAACCTCAGC 299

Qy         45 LeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPhe 64
              ::::::      :::   :::      :::||||||   ||||||||||||||||
Db        300 ATCCTGGCAGCTGTCTTGGTGGAGCCCAAACTCCACACCCCCATGTACTTCTTCCTGGGG 359

Qy         65 AsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMetSer 84
              ||||||||:::::|||             :::      :::|||   ||||||
Db        360 AACCTATCAGTGCTGGATGTTGGGTGCATCAGCGTCACTGTTCCATCAATGTTGAGTCGT 419

Qy         85 PheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePhe 104
                :::||||||||      :::            |||:::::|||   |||||||||
Db        420 CTCCTGTCCCGCAAGCGTGCAGTTCCCTGTGGGGCCTGCCTTACCCAGCTCTTCTTCTTC 479

Qy        105 CysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyrVal 124
                ||||||      :::   |||:::|||:::||||||      |||:::::::::
Db        480 CATCTGTTCGTTGGAGTGGACTGCTTCCTGCTGACCGCCATGGCCTATGACCAATTCCTG 539

Qy        125 GlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeu 144
                ||||||   ||||||   |||:::      ||||||      |||      :::|||
Db        540 GCCATCTGCCGGCCCCTCACCTACAGCACCCGCATGAGTCAGACAGTCCAGAGGATGTTG 599

Qy        145 LeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleVal 164
              :::      :::                  |||:::   ||||||      :::
Db        600 GTGGCTGCGTCCTGGGCTTGTGCTTTCACCAACGCACTGACCCACACTGTGGCCATGTCC 659

Qy        165 PheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeu 184
                |||   ||||||      |||:::::||||||:::   ||||||:::      |||
Db        660 ACGCTCAACTTCTGTGGCCCCAATGTGATCAATCACTTCTACTGTGACCTCCCACAGCTC 719

Qy        185 LeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePheIleValValThr 204
              :::||||||||:::   :::   :::|||   |||::::::|||   |||
Db        720 TTCCAGCTCTCCTGCTCCAGCACCCAACTCAATGAGCTGCTGCTTTTTGCTGTGGGTTTT 779

Qy        205 ValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSer 224
              :::      |||   |||:::      :::   |||||||||      :::   ::::::
Db        780 ATAATGGCAGGTACCCCCATGGCTCTCATTGTCATCTCCTATATCCACGTGGCAGCTGCA 839

Qy        225 IleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyr 244
              :::||||||:::   |||   |||||||||   ||||||||||||:::|||   |||:::
Db        840 GTCCTGCGAATTCGCTCTGTAGAGGGCAGGAAGAAAGCCTTCTCCACATGTGGCTCCCAC 899

Qy        245 IleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSer 264
              :::      |||::::::|||:::|||||||||   |||   |||::::::      |||
Db        900 CTCACTGTGGTTGCCATATTCTATGGTTCAGGTATCTTTAACTATATGCGACTGGGTTCA 959

Qy        265 IleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThrValValProMet 284
                        |||:::   |||      :::|||   |||   :::   ||||||
Db        960 ACCAAGCTTTCAGACAAGGATAAAGCTGTTGGAATTTTCAACACTGTCATCAATCCCATG 1019

Qy        285 PheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArg 303
              ||||||:::|||||||||   ||||||   ||||||:::   |||:::   |||
Db       1020 CTGAACCCAATCATCTACAGCTTCAGAAACCCTGATGTGCAGAGTGCCATCTGGAGG 1076

RESULT 4
US-09-668-680-10
; Sequence 10, Application US/09668680
; Patent No. 6436703
; GENERAL INFORMATION:
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION:  No. 6436703el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE:  790CIP2A
;  CURRENT APPLICATION NUMBER: US/09/668,680
;  CURRENT FILING DATE:  2000-09-22
;  PRIOR APPLICATION NUMBER:  09/649,167
;  PRIOR FILING DATE:  2000-08-23
;  PRIOR APPLICATION NUMBER:  09/540,217
```

```
;  PRIOR FILING DATE:  2000-03-31
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  pt_FL_genes Version 2.0
; SEQ ID NO 10
;    LENGTH: 1062
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (1)..(1062)
US-09-668-680-10

Alignment Scores:
Pred. No.:              4.95e-58        Length:         1062
Score:                  621.00          Matches:        128
Percent Similarity:     61.36%          Conservative:   61
Best Local Similarity:  41.56%          Mismatches:     117
Query Match:            39.38%          Indels:         2
DB:                     4               Gaps:           2

US-10-023-597-24 (1-311) x US-09-668-680-10 (1-1062)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||   |||   |||   :::          ||||||||||||   ||||||
Db        124 ATGGAAGCAGGAAACCAAACAGGATTTTTAGAGTTTATCCTTCTCGGACTCTCTGAGGAT 183

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   ||||||      :::|||   |||||||||      |||      ||||||:::
Db        184 CCAGAACTACAGCCGTTCATATTTGGGCTGTTCCTGTCCATGTACCTGGGGACGGTGCTG 243

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              |||||||||   :::|||:::   |||       :::|||   ||||||   |||||||||
Db        244 GGAAACCTGCTCATCATCCTGGCCATCAGCTCTGACTCCCACCTCCACACCCCCATGTAC 303

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||       |||||||||   ||||||       |||||||||   |||:::||||||
Db        304 TTCTTCCTCTCCAACCTGTCCTGGGTTGACATCTGTTTCAGCACTTGCATCGTCCCCAAG 363

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||::::::      :::             ||||||:::      |||:::::|||
Db        364 ATGCTGGTGAACATCCAGACCGAGAACAAAGCCATCTCCTACATGGACTGCCTCACACAG 423

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 :::|||      ||||||         ::::::   :::|||:::   |||
Db        424 GTCTATTTCTCCATGTTTTTTCCTATTCTGGACACGCTACTCCTGACCGTGATGGCCTAT 483

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||:::|||   :::|||:::||||||   |||   |||   |||:::|||   :::
Db        484 GACCGGTTTGTGGCTGTCTGCCACCCTCTGCACTATATGATCATCATGAACCCCCACCTC 543

Qy        141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||   ||||||:::         :::   :::|||||||      ::::::   |||
Db        544 TGTGGCCTCCTGGTTTTTGTCACCTGGCTCATTGGTGTCATGACATCCCTCCTCCATATT 603

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
                  ::::::   |||   ||||||   |||       :::   |||:::   |||:::
Db        604 TCTCTGATGATGCATCTAATCTTCTGTAAAGATTTTGAAATTCCACATTTTTTCTGCGAA 663

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              :::      :::|||:::|||:::|||:::   :::::::::|||      :::|||:::
Db        664 CTGACGTACATCCTCCAGCTGGCCTGCTCTGATACCTTCCTGAACAGCACGTTGATATAC 723

Qy        201 IleValValThrValGlyIleGlyVal---ProIleValAlaValPheIleSerTyrGly 219
                 :::      |||   :::||||||   |||::::::   :::      ||||||
Db        724 TTTATGACGGGTGTG---CTGGGCGTTTTTCCCCTCCTTGGGATCATTTTCTCTTATTCA 780

Qy        220 PheIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSer 239
                 |||   |||||||||   ::::::||||||:::   |||:::   ||||||   |||
Db        781 CGAATTGCTTCATCCATAAGGAAGATGTCCTCATCTGGGGGAAAACAAAAAGCACTTTCC 840

Qy        240 SerCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyr 259
              :::|||   |||::::::      ||||||||||||:::|||:::|||          :::
Db        841 ACCTGTGGGTCTCACCTCTCCGTCGTTTCTTTATTTTATGGGACAGGCATTGGGGTCCAC 900

Qy        260 LeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThr 279
                                                    :::      |||:::|||:::   ||||||
Db        901 TTCACTTCTGCGGTGACTCACTCTTCCCAGAAAATCTCCGTGGCCTCGGTGATGTACACT 960

Qy        280 ThrValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeu 299
                  |||   ||||||   ||||||   ||||||||||||||||||||||||||||
Db        961 GTGGTCACCCCCATGTTGAACCCCTTCATCTACAGCCTGAGGAACAAGGATGTGAAGGGA 1020

Qy        300 AlaLeuLysArgThrPheSerArg 307
                  ||||||             ||||||
Db       1021 GCCCTGGGGAGTCTCCTCAGCAGG 1044

RESULT 5
US-09-016-434-1312
; Sequence 1312, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1312:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1854 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g32085
US-09-016-434-1312

Alignment Scores:
Pred. No.:              1.8e-57         Length:         1854
Score:                  619.00          Matches:        131
Percent Similarity:     62.38%          Conservative:   58
Best Local Similarity:  43.23%          Mismatches:     114
Query Match:            39.25%          Indels:         0
DB:                     4               Gaps:           0

US-10-023-597-24 (1-311) x US-09-016-434-1312 (1-1854)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||      :::|||   :::|||:::::::::|||:::|||   ||||||      |||
```

```
Db        320 ATGATGGGACAAAATCAAACCAGCATCTCAGACTTCCTGCTCCTGGGCCTGCCCATCCAA 379

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal  40

Db        380 CCAGAGCAGCAAAACCTGTGCTATGCCCTGTTCTTGGCCATGTATCTTACCACCCTCCTG 439

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr  60

Db        440 GGGAACCTCCTCATCATTGTCCTCATTCGACTGGACTCCCATCTCCACACGCCTATGTAT 499

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys  80

Db        500 TTGTTTCTCAGCAACTTGTCCTTCTCTGACCTCTGCTTCTCTTCCGTGACCATTCCCAAG 559

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        560 TTGTTACAGAACATGCAGAACCAGGACCCATCCATCCCCTATGCGGACTGCCTGACCCAA 619

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        620 ATGTACTTCTTCCTGTTATTTGGAGACCTGGAGAGCTTCCTCCTTGTGGCCATGGCCTAT 679

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        680 GACCGCTATGTGGCCATCTGCTTCCCCCTGCACTACACCGCCATCATGAGCCCCATGCTC 739

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        740 TGTCTCGCCCTGGTGGCGCTGTCCTGGGTGCTGACCACCTTCCATGCCATGTTACACACT 799

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        800 TTACTCATGGCCAGGTTGTGTTTTTGTGCAGACAATGTGATCCCCCACTTTTTCTGTGAT 859

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        860 ATGTCTGCTCTGCTGAAGCTGGCCTTCTCTGACACTCGAGTTAATGAATGGGTGATATTT 919

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        920 ATCATGGGAGGGCTCATTCTTGTCATCCCATTCCTACTCATCCTTGGGTCCTATGCAAGA 979

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        980 ATTGTCTCCTCCATCCTCAAGGTCCCTTCTTCTAAGGGTATCTGCAAGGCCTTCTCTACT 1039

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db       1040 TGTGGCTCCCACCTGTCTGTGGTGTCACTGTTCTATGGAACCGTTATTGGTCTCTACTTA 1099

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db       1100 TGCTCATCAGCTAATAGTTCTACTCTAAAGGACACTGTCATGGCTATGATGTACACTGTG 1159

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db       1160 GTGACCCCCATGCTGAACCCCTTCATCTACAGCCTGAGGAACAGAGACATGAAGGGAGCC 1219

Qy        301 LeuLysArg 303

Db       1220 CTGAGCAGA 1228

RESULT 6
US-08-748-506-6
; Sequence 6, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-6

Alignment Scores:
Pred. No.:              1.51e-57        Length:         966
Score:                  616.00          Matches:        130
Percent Similarity:     59.47%          Conservative:   49
Best Local Similarity:  43.19%          Mismatches:     122
Query Match:            39.06%          Indels:         0
DB:                     3               Gaps:           0

US-10-023-597-24 (1-311) x US-08-748-506-6 (1-966)

Qy          4 GluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGlnProGlyLeu  23

Db         28 GAGAATAGTTTGTCTGTCAAACGCTTTGCATTTGCCAAGTTCTCTGAGGTCCCTGGAGAA  87

Qy         24 GlnValProValPheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeu  43

Db         88 TGCTTCCTCCTGTTCACCCTCATCCTTCTCATGTTCTTAGTATCACTAACAGGAAATGCT 147

Qy         44 GlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePhePro  63

Db        148 CTCATAGCCCTTGCTGNTTGTACCAGTCCATCTCTACACACCCCCATGTACTTCTTTCTG 207

Qy         64 PheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMet  83

Db        208 GCCAACTTGTCTCTCCTGGAGATTGGCTATACTTGCTCTGTCATACCCAAGATGCTGCAG 267

Qy         84 SerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhe 103

Db        268 AGCCTTGTGAGTGAGGCCAGAGAGATCTCTAGGGAGGGATGTGCCACACAGATGTTTTTC 327

Qy        104 PheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyr 123

Db        328 TTCACATTTTTTGGCATAACTGAGTGCTGCCTATTGGCAGCCATGGCCTTTGACCGCTGC 387

Qy        124 ValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeu 143

Db        388 ATGGGCATATGCTCCCCACTCCACTATGCAACCCGAATGAGTCGTGAGGTATGTGCCCAT 447

Qy        144 LeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIle 163

Db        448 TTGGCAATTGTTTCATGGGGAATGGGATGCATAGTAGGTCTGGGACAGACCAATTTNATT 507
```

```
Qy        164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuPro 183
              |||   ||||||             ::::::|||:::   |||||:::   |||
Db        508 TNCTCCTTGAACTTCTGTGGACCTTGTGAGATAGACCACTTCTTCTGTGACCTTCCACCT 567

Qy        184 LeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePheIleValVal 203
              ||||||   |||:::|||     :::     |||          ||||||:::
Db        568 CTCCTGGCACTTGCCTGTGGTGATACATCCCAAAACGAGGCTGCCATCTTTGTGGCAGCA 627

Qy        204 ThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPheIleLeuSer 223
                 :::   |||      |||   :::   :::      ||||||      ||||||
Db        628 ATCCTCTGTATATCTAGTCCATTTTTGGTGATCCTTTATTCTTATGTCAGAATTCTCGTT 687

Qy        224 SerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSer 243
              ::::::|||   :::   |||   |||||||||   ||||||   |||:::|||||||||
Db        688 GCAGTGCTGGTGATGCCTTCACCTGAGGGGCGCCATAAAGCTCTTTCCACCTGTTCCTCA 747

Qy        244 TyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProPro 263
              :::::::::   |||:::||||||:::||||||   :::|||||||||||||:::|||
Db        748 CACCTACTTGTAGTCACACTCTTTTATGGCTCTGTGTCCTTTACCTATTTGAGGCCCAAG 807

Qy        264 SerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThrValValPro 283
              |||                  |||:::   :::||||||||||||   |||
Db        808 TCTAGCCACTCACCAGGAATGGACAAACTCTTGGCCCTCTTCTACACAGCAGTGACATCC 867

Qy        284 MetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArg 303
              |||   ||||||:::||||||||||||||||||||||:::||||||   ||||||:::|||
Db        868 ATGCTGAACCCTATCATCTACAGTCTAAGGAACAAGGAAGTCAAGGCAGCACTGAGAAGA 927

Qy        304 Thr 304
              |||
Db        928 ACT 930

RESULT 7
US-08-748-506-5
; Sequence 5, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-5
```

```
Alignment Scores:
Pred. No.:              2.48e-57        Length:         966
Score:                  614.00          Matches:        129
Percent Similarity:     59.47%          Conservative:   50
Best Local Similarity:  42.86%          Mismatches:     122
Query Match:            38.93%          Indels:         0
DB:                     3               Gaps:           0

US-10-023-597-24 (1-311) x US-08-748-506-5 (1-966)

Qy          4 GluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGlnProGlyLeu 23
              |||||||||   :::|||   |||     |||          ||||||
Db         28 GAGAATAGTTTGACTGTCAAACACTTTGCATTTGCCAAGTTCTCTGAGGTCCCTGGAGAA 87

Qy         24 GlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeu 43
                     :::|||   |||   |||      :::   |||::::::   |||||||
Db         88 TGCTTCCTCCTGTTCAACCTCATCCTTCTCATGTTCTTAGTATCACTAACAGGAAATACT 147

Qy         44 GlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePhePro 63
                :::::::::   |||      :::      ||||||   ||||||||||||||||
Db        148 CTCATAGTCCTTGCTATTTGTACCAGTCCATCTCTACACACCCCCATGTACTTCTTTCTG 207

Qy         64 PheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMet 83
              ||||||||||||:::::      ::::::   ::::::||||||||||||||||
Db        208 GCCAACTTGTCTCTCCTGGAGATTGGCTATACTTGCTCTGTCATACCCAAGATGCTGCAG 267

Qy         84 SerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhe 103
              |||   ||||||          ||||||     ||||||   :::|||   ||||||
Db        268 AGCCTTGTGAGTGAGGCCAGAGAGATCTCTCGGGAGGGATGTGCCACACAGATGTTTTTT 327

Qy        104 PheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyr 123
              |||   ||||||       :::|||      :::|||:::||||||      ||||||
Db        328 TTCGCATTTTTTGGTATAACTGAGTGCTGCCTATTGGCAGCCATGGCCTTTGACCGCTGC 387

Qy        124 ValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeu 143
              :::   ||||||:::||||||   |||          ||||||   :::||||||
Db        388 ATGGCCATATGCTCCCCACTCCACTATGCAACCCGAATGAGTCGNGAGGTATGTGCCCAT 447

Qy        144 LeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIle 163
              |||   :::      :::|||||||||          ::::::       |||      |||
Db        448 TTGGCAATTGTTTCATGGGGAATGGGATGCATAGTAAGTCTGGGACAAACCAATTTTATT 507

Qy        164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuPro 183
              |||   ||||||             ::::::|||:::   |||||:::   |||
Db        508 TTCTCCTTGAACTTCTGTGGACCCTGTGAAATAGACCACTTCTTCTGTGACCTTCCACCT 567

Qy        184 LeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePheIleValVal 203
              ||||||   |||:::|||     :::     |||          ||||||:::|||
Db        568 CTCCTGGCACTTGCCTGTGGAGATACATCCCAAAACGAGGCTGCCATCTTTGTGGTAGCA 627

Qy        204 ThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPheIleLeuSer 223
                 :::   |||      |||   :::   :::      ||||||      ||||||
Db        628 GTCCTCTGCATATCTAGCCCATTTTTGCTGATCATTTATTCTTATGTCAAAATTCTCATT 687

Qy        224 SerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSer 243
              ::::::|||   :::   |||   |||||||||   ||||||   |||:::|||||||||
Db        688 GCAGTGCTNCTGATGCCTTCACCTGAGGGGCGCCATAAAGCTCTTTCCACCTGTTCGTCT 747

Qy        244 TyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProPro 263
              :::::::::   |||:::||||||:::||||||          ||||||||||:::|||
Db        748 CACCTACTTGTAGTCACACTTTTTTATGGCTCAGCATGTATTACCTATTTGAGGCCCAAG 807

Qy        264 SerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThrValValPro 283
              |||                  |||      :::||||||||||||   |||
Db        808 TCTAGCCACTCACCAGGAATGGACAAATTCTTGGCCCTCTTCTACACAGTAGTGACATCC 867

Qy        284 MetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArg 303
              |||   ||||||:::||||||||||||||||||||||:::||||||   ||||||:::|||
Db        868 ATGCTGAACCCTATCATCTATAGTTTAAGGAACAAGGAAGTCAAGGCAGCACTGAGAAGA 927

Qy        304 Thr 304
              |||
```

```
Db        928 ACT 930

RESULT 8
US-08-748-506-8
; Sequence 8, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-8

Alignment Scores:
Pred. No.:              1.11e-56        Length:         966
Score:                  608.00          Matches:        130
Percent Similarity:     59.33%          Conservative:   48
Best Local Similarity:  43.33%          Mismatches:     122
Query Match:            38.55%          Indels:         0
DB:                     3               Gaps:           0

US-10-023-597-24 (1-311) x US-08-748-506-8 (1-966)

Qy          4 GluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGlnProGlyLeu 23
              |||||||||   ||||||     |||      |||                 ||||||
Db         28 GAGAATAGTTTGTCTGTCAAACGTTTTGCATTTGCCAAGTTCTCTGAGGTCCCTGGAGAA 87

Qy         24 GlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeu 43
                     :::|||   |||   |||      :::   |||::::::   ||||||
Db         88 TGCTTCCTCCTGTTCACCCTCATCCTTCTCATGTTCTTAGTATCACTAACAGGAAATGCT 147

Qy         44 GlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePhePro 63
                :::   :::   |||      :::      ||||||   ||||||||||||||||
Db        148 CTCATAGCCCTTGCTATTTGTACCAGTCCATCTCTACACACCCCCATGTACTTCTTTCTG 207

Qy         64 PheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMet 83
                ||||||||||||::::::     ::::::   ::::::||||||||||||||||
Db        208 GCCAACTTGTCTCTCCTGGAGATTGGCTATACTTGCTCTGTCATACCCAAGATGCTGCAG 267

Qy         84 SerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhe 103
              |||   ||||||          |||         ||||||   :::|||   ||||||
Db        268 AGTCTTGTGAGTGAGGCCCGAGAGATCTTTCAGGTGGGATGTGCCACACAGATGTTTTTC 327

Qy        104 PheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyr 123
              |||   ||||||      :::|||      :::|||:::||||||     ||||||||
Db        328 TTCATATTCTTTGGTATAACTGAGTGCTGCCTATTGGCAGCCATGGCCTTTGACCGCTAT 387

Qy        124 ValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeu 143
              :::   ||||||:::||||||   |||      ||||||   :::||||||
Db        388 ATGGCTATATGTTCCCCACTCCACTATGCAACCCGAATGAGTCGTGAGGTATGTGCCCAC 447

Qy        144 LeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIle 163
              |||   :::      :::   ||||||            :::      |||      |||
Db        448 TTGGCAATTGTTTCATGGGTGATGGGATGCATAGTAGGTCTGGGACAGACCAATTTTATT 507

Qy        164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuPro 183
                  |||   ||||||           ::::::|||:::   |||||:::   |||
Db        508 TTCTCCTTGAACTTCTGTGGACCCTGTGAGATAGACCACTTCTTCTGTGATCTTCCACCT 567

Qy        184 LeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePheIleValVal 203
              ||||||   |||:::|||      :::                 ||||||:::||||||
Db        568 CTCCTGGCACTTGCCTGTGGTGATACATCCCAAATTGAGGCTGCCATCTTTGTGGTAGTT 627

Qy        204 ThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPheIleLeuSer 223
                  :::   |||      |||   :::   :::      ||||||      ||||||
Db        628 GTCCTCTGCATATCTAGCCCTTTTTTGCTGATCATTTATTCTTATGTCAGAATTCTCGTT 687

Qy        224 SerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSer 243
              ::::::|||   :::   |||   |||||||||   ||||||   |||:::|||||||||
Db        688 GCAGTGCTGGTGATGCCTTCACCTGAGGGGCGCCACAAAGCCCTTTCAACCTGTTCCTCC 747

Qy        244 TyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProPro 263
              :::::::::   |||:::|||||||:::|||||||||:::   |||||||||:::|||
Db        748 CACCTACTTGTAGTCACACTCTTTTATGGCTCAGGATCTGTTACCTATTTGAGGCCTAAG 807

Qy        264 SerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThrValValPro 283
              |||                ||||:::   :::||||||||||||   |||
Db        808 TCTAGCCACTCACCAGGAATGGACAAACTCTTGGCCCTCTTCTACACAGCAGTGACATCC 867

Qy        284 MetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArg 303
              |||   ||||||:::||||||||||||||||||||||||||||||   ||||||:::|||
Db        868 ATGTTGAACCCTATCATCTATAGTTTAAGGAACAAGGATGTCAAGGCAGCACTGAGAAGA 927

RESULT 9
US-09-546-986A-5
; Sequence 5, Application US/09546986A
; Patent No. 6635741
; GENERAL INFORMATION:
;  APPLICANT: Powers, Scott
;  APPLICANT:  Yang, Jianxin
;  APPLICANT:  Cutler, Gene
;  APPLICANT:  Tularik Inc.
;  TITLE OF INVENTION: No. 6635741el G-Protein Coupled Receptors
;  FILE REFERENCE: 018781-004720US
;  CURRENT APPLICATION NUMBER: US/09/546,986A
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 09/524,730
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1351
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (86)..(1108)
;   OTHER INFORMATION: human breast cancer amplified G-protein coupled
;   OTHER INFORMATION: receptor 3 (BCA-GPCR-3)
US-09-546-986A-5

Alignment Scores:
Pred. No.:              2.94e-56        Length:         1351
Score:                  606.00          Matches:        120
Percent Similarity:     62.71%          Conservative:   70
```

```
Best Local Similarity:  39.60%          Mismatches:    111
Query Match:            38.43%          Indels:        2
DB:                     4               Gaps:          2

US-10-023-597-24 (1-311) x US-09-546-986A-5 (1-1351)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||         |||   ||||||         |||:::|||   |||         :::
Db        149 ATGGAAATAGCCAATGTGAGTTCTCCAGAAGTCTTTGTCCTCCTGGGCTTCTCCGCACGA 208

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   |||:::      :::|||   :::   |||   ||||||   |||:::::::::
Db        209 CCCTCACTAGAAACTGTCCTCTTCATAGTTGTCTTGAGTTTTTACATGGTATCGATCTTG 268

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGly---LeuAsnSerArgLeuHisIleProMet 59
              ||||||   |||:::||||||||:::         :::       ||||||   ||||||
Db        269 GGCAAT---GGCATCATCATTCTGGTCTCCCATACAGATGTGCACCTCCACACACCTATG 325

Qy         60 TyrPhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIlePro 79
              |||||||||      |||||||||   :::|||   ||||||:::|||:::|||:::|||
Db        326 TACTTCTTTCTTGCCAACCTCTCCTTCCTGGACATGAGCTTCACCACGAGCATTGTCCCA 385

Qy         80 LysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSer 99
              ::::::|||   :::           :::       ||||||:::   ||||||:::
Db        386 CAGCTCCTGGCTAACCTCTGGGGACCACAGAAAACCATAAGCTATGGAGGGTGTGTGGTC 445

Qy        100 GlnPhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetVal 119
              ||||||:::         :::         :::|||      :::|||:::   |||
Db        446 CAGTTCTATATCTCCCATTGGCTGGGGGCAACCGAGTGTGTCCTGCTGGCCACCATGTCC 505

Qy        120 GluAspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGln 139
              |||||||||      ||||||   ||||||   ||||||:::   |||   ||||||
Db        506 TATGACCGCTACGCTGCCATCTGCAGGCCACTCCATTACACTGTCATTATGCATCCACAG 565

Qy        140 ValCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHis 159
              :::||||||   |||   |||       :::       |||:::      ::::::
Db        566 CTTTGCCTTGGGCTAGCTTTGGCCTCCTGGCTGGGGGGTCTGACCACCAGCATGGTGGGC 625

Qy        160 ThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCys 179
              :::       :::   |||       |||   :::|||   ::::::|||:::   |||
Db        626 TCCACGCTCACCATGCTCCTACCGCTGTGTGGGAACAATTGCATCGACCACTTCTTTTGC 685

Qy        180 AspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIle 199
              ::::::      :::::::::|||:::|||     :::   :::|||   :::   :::
Db        686 GAGATGCCCCTCATTATGCAACTGGCTTGTGTGGATACCAGCCTCAATGAGATGGAGATG 745

Qy        200 PheIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGly 219
              ::::::      |||   :::   :::|||:::      :::   :::|||||||||
Db        746 TACCTGGCCAGCTTTGTCTTTGTTGTCCTGCCTCTGGGGCTCATCCTGGTCTCTTACGGC 805

Qy        220 PheIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSer 239
              |||       ::::::|||::::::   ||||||||||||||   |||||||||:::
Db        806 CACATTGCCCGGGCCGTGTTGAAGATCAGGTCAGCAGAAGGGCGGAGAAAGGCATTCAAC 865

Qy        240 SerCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyr 259
              :::|||||||||::::::      |||||||||||:::||||||       |||   |||
Db        866 ACCTGTTCTTCCCACGTGGCTGTGGTGTCTCTGTTTTACGGGAGCATCATCTTCATGTAT 925

Qy        260 LeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThr 279
              |||:::|||                  :::||||||||||      :::||||||||||||
Db        926 CTCCAGCCAGCCAAGAGCACCTCCCATGAGCAGGGCAAGTTCATAGCTCTGTTCTACACC 985

Qy        280 ThrValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeu 299
              |||   |||      ||||||||||||||:::|||||||||   :::||||||
Db        986 GTAGTCACTCCTGCGTTGAACCCACTTATTTACACCCTGAGGAACACGGAGGTGAAGAGC 1045

Qy        300 AlaLeuLys 302
              ||||||:::
Db       1046 GCCCTCCGG 1054

RESULT 10
US-09-524-730-5
; Sequence 5, Application US/09524730
; Patent No. 6638733
; GENERAL INFORMATION:
;  APPLICANT:  Powers, Scott
;  APPLICANT:  Yang, Jianxin
;  APPLICANT:  Cutler, Gene
;  APPLICANT:  Tularik Inc.
;  TITLE OF INVENTION:  No. 6638733el G-Protein Coupled Receptors
;  FILE REFERENCE:  018781-004710US
;  CURRENT APPLICATION NUMBER: US/09/524,730
;  CURRENT FILING DATE:  2000-03-14
;  NUMBER OF SEQ ID NOS:  16
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1351
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (86)..(1108)
;   OTHER INFORMATION: human breast cancer amplified G-protein coupled
;   OTHER INFORMATION: receptor 3 (BCA-GPCR-3)
US-09-524-730-5

Alignment Scores:
Pred. No.:              2.94e-56        Length:         1351
Score:                  606.00          Matches:        120
Percent Similarity:     62.71%          Conservative:   70
Best Local Similarity:  39.60%          Mismatches:     111
Query Match:            38.43%          Indels:         2
DB:                     4               Gaps:           2

US-10-023-597-24 (1-311) x US-09-524-730-5 (1-1351)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||         |||   ||||||         |||:::|||   |||         :::
Db        149 ATGGAAATAGCCAATGTGAGTTCTCCAGAAGTCTTTGTCCTCCTGGGCTTCTCCGCACGA 208

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   |||:::      :::|||   :::   |||   ||||||   |||:::::::::
Db        209 CCCTCACTAGAAACTGTCCTCTTCATAGTTGTCTTGAGTTTTTACATGGTATCGATCTTG 268

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGly---LeuAsnSerArgLeuHisIleProMet 59
              ||||||   |||:::||||||||:::         :::       ||||||   ||||||
Db        269 GGCAAT---GGCATCATCATTCTGGTCTCCCATACAGATGTGCACCTCCACACACCTATG 325

Qy         60 TyrPhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIlePro 79
              |||||||||      |||||||||   :::|||   ||||||:::|||:::|||:::|||
Db        326 TACTTCTTTCTTGCCAACCTCTCCTTCCTGGACATGAGCTTCACCACGAGCATTGTCCCA 385

Qy         80 LysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSer 99
              ::::::|||   :::           :::       ||||||:::   ||||||:::
Db        386 CAGCTCCTGGCTAACCTCTGGGGACCACAGAAAACCATAAGCTATGGAGGGTGTGTGGTC 445

Qy        100 GlnPhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetVal 119
              ||||||:::         :::         :::|||      :::|||:::   |||
Db        446 CAGTTCTATATCTCCCATTGGCTGGGGGCAACCGAGTGTGTCCTGCTGGCCACCATGTCC 505

Qy        120 GluAspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGln 139
              |||||||||      ||||||   ||||||   ||||||:::   |||   ||||||
Db        506 TATGACCGCTACGCTGCCATCTGCAGGCCACTCCATTACACTGTCATTATGCATCCACAG 565

Qy        140 ValCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHis 159
              :::||||||   |||   |||       :::       |||:::      ::::::
Db        566 CTTTGCCTTGGGCTAGCTTTGGCCTCCTGGCTGGGGGGTCTGACCACCAGCATGGTGGGC 625

Qy        160 ThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCys 179
              :::       :::   |||       |||   :::|||   ::::::|||:::   |||
Db        626 TCCACGCTCACCATGCTCCTACCGCTGTGTGGGAACAATTGCATCGACCACTTCTTTTGC 685

Qy        180 AspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIle 199
              ::::::      :::::::::|||:::|||     :::   :::|||   :::   :::
Db        686 GAGATGCCCCTCATTATGCAACTGGCTTGTGTGGATACCAGCCTCAATGAGATGGAGATG 745
```

```
Qy        200 PheIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGly 219
              ::::::          |||   :::   :::|||:::      :::   :::||||||||
Db        746 TACCTGGCCAGCTTTGTCTTTGTTGTCCTGCCTCTGGGGCTCATCCTGGTCTCTTACGGC 805

Qy        220 PheIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSer 239
                |||      ::::::|||::::::   ||||||||||||||   |||||||||:::
Db        806 CACATTGCCCGGGCCGTGTTGAAGATCAGGTCAGCAGAAGGGCGGAGAAAGGCATTCAAC 865

Qy        240 SerCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyr 259
              :::|||||||||::::::     ||||||||||||:::||||||      |||   |||
Db        866 ACCTGTTCTTCCCACGTGGCTGTGGTGTCTCTGTTTTACGGGAGCATCATCTTCATGTAT 925

Qy        260 LeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThr 279
              |||:::|||                  :::||||||||||      :::||||||||||||
Db        926 CTCCAGCCAGCCAAGAGCACCTCCCATGAGCAGGGCAAGTTCATAGCTCTGTTCTACACC 985

Qy        280 ThrValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeu 299
                |||   |||       ||||||||||||||:::||||||||   :::||||||
Db        986 GTAGTCACTCCTGCGTTGAACCCACTTATTTACACCCTGAGGAACACGGAGGTGAAGAGC 1045

Qy        300 AlaLeuLys 302
              ||||||:::
Db       1046 GCCCTCCGG 1054

RESULT 11
US-08-748-506-7
; Sequence 7, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-7

Alignment Scores:
Pred. No.:              3.64e-55        Length:         966
Score:                  594.00          Matches:        129
Percent Similarity:     59.14%          Conservative:   49
Best Local Similarity:  42.86%          Mismatches:     123
Query Match:            37.67%          Indels:         0
DB:                     3               Gaps:           0
```

```
US-10-023-597-24 (1-311) x US-08-748-506-7 (1-966)

Qy          4 GluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGlnProGlyLeu 23
              |||||||||   ||||||      |||       |||                ||||||
Db         28 GAGAATAGTTTGTCTGTCAAACGCTTTGCATTTGCCAAGTTCTCTGAGGTCCCTGGAGAA 87

Qy         24 GlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeu 43
                    :::|||   |||   |||       :::   |||::::::   ||||||
Db         88 TGCTTCCTCCTGTTCACCCTCATCCTTCTCATGTTCTTAGTATCACTAACAGGAAATACT 147

Qy         44 GlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePhePro 63
                 :::   :::   |||       :::          ||||||   ||||||||||||||||
Db        148 CTCATAGCCCTTGCTATTTGTACCAGTCCATCTCTACACACCCCCATGTACTTCTTTCTG 207

Qy         64 PheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMet 83
              ||||||||||||:::::      ::::::   ::::::||||||||||||||||
Db        208 GCCAACTTGTCTCTCCTGGAGATTGGCTATACTTGCTCTGTCATACCCAAGATGCTGCAG 267

Qy         84 SerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhe 103
              |||   ||||||                ||||||:::   ||||||   ||||||   ||||||
Db        268 AGCCTTGTGAGTGAGGCCCGAGGGATCTCTTGGGAGGGTTGTGCCTCACAGATGTTCTTC 327

Qy        104 PheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyr 123
              |||   ||||||      :::|||      :::|||:::||||||      |||||||||
Db        328 TTCATATTCTTTGGTATAACTGAGTGCTGCCTATTGGCAGCCATGGCCTTTGACCGCTAT 387

Qy        124 ValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeu 143
              :::   ||||||:::||||||   |||       ||||||      ||||||
Db        388 ATGGCTATATGTTCCCCACTCCACTATGCAACCCGAATGAGTCGTGGGGTATGTGCCTAT 447

Qy        144 LeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIle 163
              |||   :::      :::   ||||||            :::      |||      |||
Db        448 TTGGCAATTGTCTCATGGGTGATGGGATGCATAGTAGGTCTGGGACAGACCAATTTTATT 507

Qy        164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuPro 183
                    |||   ||||||            ::::::|||:::   ||||||:::   |||
Db        508 TTCTCCTTGAACTTCTGTGGACCCTGTGAGATAGACCACTTCTTCTGTGACCTTCCACCT 567

Qy        184 LeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePheIleValVal 203
              ||||||   |||:::|||      :::      |||          ||||||:::
Db        568 CTCCTGGCACTTGCCTGTGGTGATACATCCCAAAATGAGGCTGCCATCTTTGTGGCAGCA 627

Qy        204 ThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPheIleLeuSer 223
              :::   |||      |||   :::   :::          ||||||      ||||||
Db        628 GTGCTCTGCATATTTAGTCCATTTTTACTGATCATTTCTTCCTATGTCAGAATTCTCGTT 687

Qy        224 SerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSer 243
              ::::::|||   :::   |||   |||||||||   ||||||   |||:::|||||||||
Db        688 GCAGTGCTGGTGATGCCTTCACCTGAGGGGCGCCATAAAGCTCTCTCTACCTGTTCATCT 747

Qy        244 TyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProPro 263
              :::::::::   |||:::||||||:::||||||   :::   |||||||||:::
Db        748 CACCTACTTGTAGTCACACTCTTCTATGGCTCAACATCTGCCACCTATTTGAGGTCCAAG 807

Qy        264 SerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThrValValPro 283
              |||                  ||:::   :::||||||||||||:::|||
Db        808 TCTAGCCACTCACCAGGAGTGGACAAACTCTTGGCCCTCTTCTATACATCAGTGACATCC 867

Qy        284 MetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArg 303
              |||   ||||||:::||||||||||||||||||||:::||||||   ||||||:::|||
Db        868 ATGCTGAATCCCATCATCTACAGCTTAAGGAACAAGGAAGTAAAGGGTGCACTGAGAAGA 927

Qy        304 Thr 304
              |||
Db        928 ACT 930

RESULT 12
US-09-016-434-1363
; Sequence 1363, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
```

```
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1363:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3459 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g425220
US-09-016-434-1363

Alignment Scores:
Pred. No.:              4.15e-54        Length:         3459
Score:                  591.50          Matches:        125
Percent Similarity:     58.63%          Conservative:   55
Best Local Similarity:  40.72%          Mismatches:     126
Query Match:            37.51%          Indels:         1
DB:                     4               Gaps:           1

US-10-023-597-24 (1-311) x US-09-016-434-1363 (1-3459)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||         |||   |||      :::||||||:::|||   |||:::
Db        883 ATGGATGGAGGCAACCAGAGTGAAGGTTCAGAGTTCCTTCTCCTGGGGATGTCAGAGAGT 942

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||       |||       :::|||::::::||||||      |||   ||||||||||||
Db        943 CCTGAGCAGCAGCAGATCCTGTTTTGGATGTTCCTGTCCATGTACCTGGTCACGGTGGTG 1002

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||:::   :::|||:::   |||      :::||||||||||||   |||:::|||
Db       1003 GGAAATGTGCTCATCATCCTGGCCATCAGCTCTGATTCCCGCCTGCACACCCCCGTGTAC 1062

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||      |||||||||      |||      |||   |||      |||||||||
Db       1063 TTCTTCCTGGCCAACCTCTCCTTCACTGACCTCTTCTTTGTCACCAACACAATCCCCAAG 1122

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||::::::      |||             ||||||:::   |||||:::::::|||
Db       1123 ATGCTGGTGAACCTCCAGTCCCATAACAAAGCCATCTCCTATGCAGGGTGTCTGACGCAG 1182

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 :::|||           |||      ::::::   ||||||:::   |||
Db       1183 CTCTACTTCCTGGTCTCCTTGGTGGCCCTGGACAACCTCATCCTGGCTGTGATGGCATAT 1242

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||   ||||||   ||||||      |||||||||::::::
Db       1243 GACCGCTATGTGGCCATCTGCTGCCCCCTCCACTACACCACAGCCATGAGCCCTAAGCTC 1302

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||:::|||||||||      :::   :::   |||      :::   ||||||
Db       1303 TGTATCTTACTCCTTTCCTTGTGTTGGGTCCTATCCGTCCTCTATGGCCTCATACACACC 1362

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
                    :::      :::|||||||||          :::::::::      |||:::
Db       1363 CTCCTCATGACCAGAGTGACCTTCTGTGGGTCACGAAAAATCCACTACATCTTCTGTGAG 1422

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              :::      ||||||   ::::::|||:::          ||||||      |||:::
Db       1423 ATGTATGTATTGCTGAGGATGGCATGTTCCAACATTCAGATTAATCACACAGTGCTGATT 1482

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
                                          :::|||          |||   |||||||||
Db       1483 GCCACAGGCTGCTTCATCTTCCTCATTCCCTTTGGATTCGTGATCATTTCCTATGTGCTG 1542

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              |||:::   :::|||||||||:::   |||          :::   ||||||||||:::
Db       1543 ATTATCAGAGCCATCCTCAGAATACCCTCAGTCTCTAAGAAATACAAAGCCTTCTCCACC 1602

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||:::|||::::::   ||||||||||||||:::|||:::          ||||||
Db       1603 TGTGCCTCCCATTTGGGTGCAGTCTCCCTCTTCTATGGGACACTTTGTATGGTATACCTA 1662

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||                :::   :::      |||:::::::::   |||
Db       1663 AAGCCCCTCCATACCTACTCTGTG---AAGGACTCAGTAGCCACAGTGATGTATGCTGTG 1719

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              |||   ||||||   ||||||   ||||||||||||||||||||||:::      |||
Db       1720 GTGACACCCATGATGAATCCCTTCATCTACAGCCTGAGGAACAAGGACATGCATGGGGCT 1779

Qy        301 LeuLysArgThrPheSerArg 307
              |||  |||         :::
Db       1780 CTGGGAAGACTCCTAGATAAA 1800

RESULT 13
US-08-827-291A-1
; Sequence 1, Application US/08827291A
; Patent No. 5874243
;  GENERAL INFORMATION:
;    APPLICANT:  Macina, Roberto
;    APPLICANT:  Sathe, Ganesh
;    TITLE OF INVENTION:  NOVEL OLRCC15 RECEPTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,291A
;      FILING DATE:  28-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
```

```
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GP50001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5015
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1290 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-827-291A-1

Alignment Scores:
Pred. No.:              1.91e-54        Length:         1290
Score:                  589.00          Matches:        118
Percent Similarity:     58.90%          Conservative:   64
Best Local Similarity:  38.19%          Mismatches:     123
Query Match:            37.35%          Indels:         4
DB:                     2               Gaps:           1

US-10-023-597-24 (1-311) x US-08-827-291A-1 (1-1290)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||  |||||||||:::      ::::::||||||      |||:::   :::
Db        296 ATGGCAAGGGAGAATTCGACCTTCAACTCCGACTTCATCTTCCTGGGAATCTTCAATCAC 355

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
                          :::||||||||||   |||      ::::::|||      :::
Db        356 AGCCCCACCCACACCTTCCTCTTCTTTCTGGTCCTGGCCATCTTTTCAGTGGCCTTCATG 415

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||      :::::::::||||||   |||:::::::::||||||   |||||||||
Db        416 GGAAACTCTGTCATGGTTCTCCTCATCTACCTGGACACCCAGCTCCACACCCCCATGTAC 475

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
                        |||||||||:::|||      |||   :::||||||
Db        476 CTCCTCCTCAGCCAACTGTCCCTCATGGACCTCATGCTCATCTGCACCACCGTACCCAAG 535

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              |||'      :::::::::|||          ||||||      ||||||   :::|||
Db        536 ATGGCCTTCAACTACCTGTCTGGCAGCAAGTCCATTTCTATGGCTGGTTGTGCCACACAA 595

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                  ||||||:::          :::   ||||||   |||:::|||:::   |||
Db        596 ATTTTCTTCTATACATCACTGCTTGGCTCTGAATGCTTTCTTTTGGCTGTTATGGCTTAT 655

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              |||||||||      ||||||:::||||||   ||||||      |||||||||::::::
Db        656 GACCGCTACACTGCCATTTGCCACCCTCTAAGATACACCAATCTCATGAGCCCTAAAATT 715

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||   |||:::          :::  :::|||                :::  :::
Db        716 TGTGGACTTATGACTGCCTTTTCCTGGATCCTGGGCTCTACAGATGGAATCATTTATGCT 775

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
                          ::::::|||          :::   |||:::   |||:::
Db        776 GTAGCCACATTTTCCTTCTCCTACTGTGGGTCTCGGGAAATAGCCCACTTCTTCTGTGAG 835

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              :::      ||||||  ||||||||||||   :::   |||         |||||||||
Db        836 TTACCTTCCCTACTAATCCTCTCATGCAATGACACATCAATATTTGAAAAGGTTATTTTC 895

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||         |||   :::      |||:::      :::      ||||||
Db        896 ATTTGCTCTATAGTAATGCTTGTTTTCCCTGTTGCAATCATCATTGCTTCCTATGCTGGA 955

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ::::::  :::::::::  :::   |||   |||||||||   |||||||||::::::
Db        956 GTTATTCTGGCTGTCATTCACATGGGATCTGGAGAGGGTCGTCGCAAAGCTTTCACGACC 1015

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::::::::   |||   :::|||:::|||:::|||   |||   |||:::
Db       1016 TGTTCCTCTCACCTCATGGTGGTGGGAATGTTCTATGGAGCAGGTTTGTTCATGTACATA 1075

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              :::|||   |||                |||   |||:::   |||:::|||||||||
Db       1076 CAGCCCACATCTGATCGCTCCCCAACGCAGGACAAGCTGGTGTCTGTATTCTACACCATC 1135

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::   ||||||   ||||||||||||||||||||||||||||||:::|||
Db       1136 CTCACTCCCATGCTGAATCCCCTCATCTACAGCCTCCGCAACAAGGAAGTG--------- 1186

Qy        301 LeuLysArgThrPheSerArgIleSer 309
                    |||   |||   :::||||||
Db       1187 ---ACCAGAGCATTCATGAAGATCTCA 1210

RESULT 14
US-09-016-434-1313
; Sequence 1313, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1313:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1438 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g32092
US-09-016-434-1313

Alignment Scores:
Pred. No.:              6.85e-54        Length:         1438
Score:                  584.50          Matches:        115
Percent Similarity:     61.39%          Conservative:   71
Best Local Similarity:  37.95%          Mismatches:     116
Query Match:            37.06%          Indels:         1
DB:                     4               Gaps:           1

US-10-023-597-24 (1-311) x US-09-016-434-1313 (1-1438)
```

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||       ||||||   :::   :::|||:::|||:::       |||
Db        362 ATGAAAAGAGAGAACTTTACTCTCATCACTGACTTTGTTTTCCAAGGTTTCTCTAGCTTC 421

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
                 |||:::   :::|||   :::||||||       |||   :::|||:::
Db        422 CATGAGCAGCAGATCACCCTTTTTGGCGTGTTCCTTGCACTATACATCTTAACCTTAGCA 481

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||:::   ::::::   :::|||   ::::::       ||||||   |||||||||
Db        482 GGCAATATCATCATTGTGACCATCATCCGAATTGATCTTCATCTTCACACACCCATGTAC 541

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||       ||||||      :::      ::::::      |||:::|||:::
Db        542 TTCTTCCTGAGCATGCTGTCCACTTCAGAGACTGTATATACATTGGTCATTCTCCCAAGA 601

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||   |||   |||                  :::|||       ||||||   :::|||
Db        602 ATGCTCTCCAGCCTCGTAGGTATGAGCCAGCCCATGTCATTGGCAGGGTGTGCCACACAG 661

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              |||||||||       |||       :::       |||:::|||:::||||||
Db        662 ATGTTCTTTTTTGTAACCTTTGGCATCACTAACTGCTTCCTGCTCACAGCAATGGGATAT 721

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||   |||   :::   |||:::   ::::::
Db        722 GACCGCTATGTGGCCATCTGCAACCCCCTGAGATACATGGTTATTATGAACAAGAGGCTG 781

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
                 :::   |||:::||||||          :::|||:::       |||:::
Db        782 CGTATCCAACTTGTCCTGGGGGCCTGCAGCATTGGGCTGATTGTAGCAATAACGCAAGTG 841

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
                 ::::::       |||   |||||||||       |||   |||:::   ||||||
Db        842 ACATCTGTATTCAGGTTACCCTTCTGTGCT---AGAAAGGTGCCCCACTTCTTCTGTGAC 898

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||   |||:::::::::|||||||||       :::   :::|||   ::::::
Db        899 ATCCGCCCTGTGATGAAGCTCTCCTGCATTGACACCACTGTCAATGAAATCCTGACTTTG 958

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||:::       :::   :::   ||||||:::       ||||||||||||||||
Db        959 ATTATCAGTGTGCTGGTGCTTGTTGTACCTATGGGTCTGGTTTTCATTTCTTATGTTCTC 1018

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              |||:::|||:::||||||::::::::|||   |||||||||   |||||||||::::::
Db       1019 ATTATCTCTACAATCCTCAAGATTGCTTCAGTTGAGGGCCGGAAGAAGGCTTTTGCCACC 1078

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||:::|||::::::       |||   :::   :::           :::       ||||||
Db       1079 TGTGCATCCCACCTCACTGTGGTCATTGTCCACTACAGCTGTGCCTCCATTGCCTACCTC 1138

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   |||           :::       ::::::   |||:::   ||||||
Db       1139 AAGCCCAAGTCAGAGAACACCAGAGAACATGACCAGCTGATCTCGGTGACCTACACTGTC 1198

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::   |||:::   ||||||::::::|||:::|||||||||||||:::||||||   |||
Db       1199 ATCACTCCCCTACTGAACCCTGTGGTATACACCCTGAGAAATAAAGAGGTCAAAGATGCT 1258

Qy        301 LeuLysArg 303
              |||   |||
Db       1259 CTGTGCAGG 1267
```

RESULT 15
US-09-546-986A-1
; Sequence 1, Application US/09546986A
; Patent No. 6635741
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Yang, Jianxin
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. 6635741el G-Protein Coupled Receptors
; FILE REFERENCE: 018781-004720US
; CURRENT APPLICATION NUMBER: US/09/546,986A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 09/524,730
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(974)
; OTHER INFORMATION: human breast cancer amplified G-protein coupled
; OTHER INFORMATION: receptor 1 (BCA-GPCR-1)
US-09-546-986A-1

Alignment Scores:

Pred. No.:	5.9e-53	Length:	1035
Score:	574.00	Matches:	115
Percent Similarity:	62.33%	Conservative:	72
Best Local Similarity:	38.33%	Mismatches:	107
Query Match:	36.40%	Indels:	6
DB:	4	Gaps:	3

US-10-023-597-24 (1-311) x US-09-546-986A-1 (1-1035)

```
Qy         12 PheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGlnValProValPhePheLeuPhe 31
              |||||||||   |||:::       :::|||   |||:::::|||:::|||   :::
Db         66 TTCATCCTTCTGGGTGTGTCTGACAGGCCGTGGCTGGAACTCCCTCTCTTTGTGGTCCTC 125

Qy         32 LeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIleLeuIleGlyLeu 51
              |||       |||   :::   ::::::||||||:::   :::|||:::         :::
Db        126 CTGCTGTCCTATGTGCTGGCCATGTTGGGGAACGTCGCCATCATCCTGGCATCCCGGGTG 185

Qy         52 AsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSerLeuValAspPhe 71
              :::   :::||||||   |||||||||   |||       :::||||||   :::|||
Db        186 GATCCTCAACTCCACAGCCCCATGTACATCTTCCTCAGTCACCTGTCCTTCCTGGACCTC 245

Qy         72 SerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLysAsnIle 91
                 ::::::||||||   :::|||:::||||||::::::       |||   :::
Db        246 TGCTACACCACCACGACAGTCCCTCAGATGCTGGTCAACATGGGCAGTTCCCAGAAGACC 305

Qy         92 IleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPheSerGlu 111
              |||||||:::   ||||||      |||:::      |||   :::           :::|||
Db        306 ATCAGCTATGGAGGCTGCACTGTGCAATATGCAGTCTTCCACTGGCTGGGATGCACGGAG 365

Qy        112 SerPheIleLeuSerAlaMetValGluAspArgTyrValGlyIleCysAsnProLeuLeu 131
                    :::|||:::|||||||           ||||||||||||       |||   ||||||
Db        366 TGCATCGTCCTGGCCGCCATGGCCCTGGACCGCTACGTGGCCAGCTGCAAGCCCCTGCAC 425

Qy        132 TyrThrIleThrMetSerProGlnValCys------LeuLeuLeuLeuLeuGlyValTyr 149
              |||   :::   |||          :::|||       |||:::   |||       :::
Db        426 TATGCCGTTCTCATGCACCGTGCTCTCTGTCAGCAGCTCGTGGCTCTGGCCTGGCTCAGT 485

Qy        150 GlyMetGly---ValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPhe 168
              |||   |||       |||       |||       |||       |||   |||   |||
Db        486 GGCTTCGGCAACTCCTTCGTGCAGGTGGTCCTGACG---------GTGCAATTGCCATTC 536

Qy        169 CysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSer 188
              |||       ::::::|||::::::   |||::::::       :::::::::||||||
Db        537 TGCGGGCGGCAGGTGCTGAACAACTTTTTCTGTGAGGTGCCGGCCGTGATCAAGCTGTCG 596

Qy        189 CysAsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGly 208
              |||       :::   :::|||       ::::::   ::::::|||          :::
Db        597 TGTGCTGACACCGCTATGAATGACACCATACTGGCTGTGCTGGTGGCCTTCTTCGTGTTG 656

Qy        209 ValProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgVal 228
              ||||||:::       :::   :::|||||||||||||||       ::::::||||||:::
Db        657 GTGCCCCTGGCTCTCATCCTTCTCTCCTATGGCTTTATTGCCCGGGCAGTGCTCAGGATC 716
```

```
Qy        229 SerSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaVal 248
              |||::::::||||||   |||||||||   :::||||||||::::::::::   |||
Db        717 CAGTCCTCCAAGGGACGACACAAGGCCTTTGGGACGTGTTCCTCCCACCTGATGATCGTC 776

Qy        249 SerLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeu 268
              |||||||||:::           :::   ||||||:::|||||||||
Db        777 TCCCTCTTCTACCTACCTGCGATTTACATGTATCTGCAGCCCCCTTCCAGCTACTCCCAA 836

Qy        269 AspGlnGlyLysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeu 288
              :::|||||||||      ||||||||||||:::   :::   |||      ||||||
Db        837 GAGCAGGGCAAATTTATTTCTCTCTTCTATTCCATAATCACCCCCACTCTCAATCCCTTC 896

Qy        289 IleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIle 308
              |||:::|||||||||||||||:::|||   ||||||:::|||      :::|||||||
Db        897 ACCTACACCCTGAGAAATAAAGATATGAAGGGGGCTCTGAGGAGACTTCTGGCCAGGATC 956
```

Search completed: September 30, 2004, 12:13:11
Job time : 94 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 09:21:50 ; Search time 2904 Seconds
(without alignments)
3198.050 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 1577
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163410_14892/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023597_@CGN_1_1_3437_@runat_29092004_163410_14892 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1273.5	80.8	933	29	AY407109	AY407109 Homo sapi
2	1223.5	77.6	933	29	AY407111	AY407111 Mus muscu
3	1158	73.4	903	29	AY405611	AY405611 Mus muscu
4	1139	72.2	903	29	AY405609	AY405609 Homo sapi
5	1086.5	68.9	799	29	AY407110	AY407110 Pan trogl
6	1024.5	65.0	930	29	AY405659	AY405659 Mus muscu
7	958.5	60.8	930	29	AY405657	AY405657 Homo sapi
8	935	59.3	899	29	AY405610	AY405610 Pan trogl
9	923.5	58.6	930	29	AY419057	AY419057 Mus muscu
10	888.5	56.3	930	29	AY419055	AY419055 Homo sapi
11	880.5	55.8	930	29	AY419056	AY419056 Pan trogl
12	812	51.5	726	28	AZ255734	AZ255734 RPCI-23-1
13	791	50.2	748	14	CB174479	CB174479 OR_2oE1H1
14	786	49.8	933	29	AY405614	AY405614 Mus muscu
15	775	49.1	933	29	AY405612	AY405612 Homo sapi
16	757.5	48.0	848	28	BZ281255	BZ281255 CH230-344
17	724	45.9	945	29	AY402815	AY402815 Mus muscu
18	710.5	45.1	625	29	CE158635	CE158635 tigr-gss-
c 19	707	44.8	733	28	BZ201201	BZ201201 CH230-303
20	706	44.8	636	14	CB174480	CB174480 OR_2oE1H1
21	701.5	44.5	942	29	AY402814	AY402814 Pan trogl
22	700	44.4	637	14	CB174340	CB174340 OR_2o42A0
23	699.5	44.4	945	29	AY402813	AY402813 Homo sapi
24	686.5	43.5	664	14	CB173618	CB173618 OR_2o21B1
25	684	43.4	626	29	CE184619	CE184619 tigr-gss-
26	679	43.1	924	29	AY421166	AY421166 Mus muscu
27	675.5	42.8	752	28	BZ102526	BZ102526 CH230-238
28	669	42.4	897	29	AY405605	AY405605 Mus muscu
29	664.5	42.1	630	28	AQ503914	AQ503914 RPCI-11-2
30	657	41.7	936	29	AY405000	AY405000 Homo sapi
31	655	41.5	936	29	AY405001	AY405001 Pan trogl
32	654	41.5	939	29	AY405002	AY405002 Mus muscu
33	651	41.3	897	29	AY405603	AY405603 Homo sapi
34	650	41.2	939	29	AY405723	AY405723 Homo sapi
35	650	41.2	945	29	AY404165	AY404165 Mus muscu
36	649	41.2	907	28	BZ243737	BZ243737 CH230-469
37	648.5	41.1	921	29	AY421164	AY421164 Homo sapi
38	648	41.1	616	28	BH342053	BH342053 CH230-51L
39	642	40.7	942	29	AY415251	AY415251 Homo sapi
40	641	40.6	731	28	BZ207385	BZ207385 CH230-487
41	641	40.6	945	29	AY415715	AY415715 Mus muscu
42	641	40.6	2021	11	BC016940	BC016940 Homo sapi
43	640	40.6	1680	11	BC051250	BC051250 Mus muscu
44	635	40.3	933	29	AY402266	AY402266 Mus muscu
45	630	39.9	948	29	AY401403	AY401403 Homo sapi

ALIGNMENTS

RESULT 1
AY407109
LOCUS AY407109 933 bp DNA linear GSS 15-DEC-2003
DEFINITION Homo sapiens OR8B12 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY407109
VERSION AY407109.1 GI:39763080
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 933)

```
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>933
                     /gene="OR8B12"
                     /locus_tag="HCM2781"
ORIGIN

Alignment Scores:
Pred. No.:              1.27e-98        Length:         933
Score:                  1273.50         Matches:        250
Percent Similarity:     87.46%          Conservative:   22
Best Local Similarity:  80.39%          Mismatches:     38
Query Match:            80.75%          Indels:         1
DB:                     29              Gaps:           1

US-10-023-597-24 (1-311) x AY407109 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||::::::|||:::|||||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACACGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTCTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||:::||||||   ||||||||||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAATAGATTTCTGTTTCTCCACTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                |||||||||||||||||||   |||||||||||||||||||||||||||||
Db        298 CTCTTCTTCTTCTGCTTCTTTGTCGTCTCTGAGTCCTTCATCCTGTCAGCGATGGCGTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   |||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTACGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||   ||||||||||||      ||||||:::||||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   |||||||||||||||||||||||||||||||:::||||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::||||||||||   ||||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||      ||||||   |||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   ||||||||||||||:::||||||||||||||||||||||||||||||||
Db        778 AAACCCCTTTCCATCCTGCCCCTCGAGCAAGGGAAAGTGTCCTCCCTGTTCTATACCATA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              :::||||||:::   ||||||||||||||||||||||||||||||||||||||:::|||
Db        838 ATAGTCCCCGTGTTAAACCCATTAATCTATAGCTTGAGGAACAAGGATGTCAAAGTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||:::||||||       |||       ||||||
Db        898 CTGAGGAGAACTTTGGGCAGAAAAATCTTTTCT 930

RESULT 2
AY407111
LOCUS       AY407111                 933 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Mus musculus OR8B12 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY407111
VERSION     AY407111.1  GI:39763082
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>933
                     /gene="OR8B12"
                     /locus_tag="HCM2781"
ORIGIN

Alignment Scores:
Pred. No.:              2.3e-94         Length:         933
Score:                  1223.50         Matches:        237
```

```
Percent Similarity:     86.97%          Conservative:  30
Best Local Similarity:  77.20%          Mismatches:    39
Query Match:            77.58%          Indels:        1
DB:                     29              Gaps:          1

US-10-023-597-24 (1-311) x AY407111 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||   |||:::|||   |||||||||   ||||||||||||||||||||||       |||
Db          1 ATGACTGCCAAGAAT---TCCTCTGTGATAGAGTTCATCCTCGCAGGCCTGACAGACCAA 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||:::::|||:::|||||||||||||||||||||||   ||||||||||||
Db         58 CCAGGACTCCGCATGCCCCTCTTCTTCTTGTTTCTAGGTTTCTACATGGTGACTGTGGTT 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||||||   ||||||   |||||||||
Db        118 GGGAACCTGGGTTTGATCTCCCTGATTGGGCTGAACTCTCACCTTCATACCCCTATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||:::::||||||   :::|||:::|||||||||||||||
Db        178 TTCTTTCTCTTCAATCTTTCTGTAATAGATTTCTGTTATTCCTCCACTATCATCCCAAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              |||||||||:::|||:::|||||||||||||||||||||||   :::|||||||||:::|||
Db        238 ATGCTTATGAATTTTATCTCAAGGAAGAACATCATCTCACACTCAGGGTGCATGACACAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||   |||||||||||||||||||||||||||
Db        298 TTGTTTTTCTTCTGTTTCTTTGTTGTCTCTGAGTCCTTCATCCTGTCAGCCATGGCATAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||||||||||:::|||||||:::|||||||||||||||||||
Db        358 GACCGTTATGTTGCCATCTGTAACCCCCTGATGTACACAGTCACCATGTCTCCTCAGGTA 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||   ||||||   |||||||:::|||||||||
Db        418 TATTTACTTCTTTTGCTGGGAGTCTATGTGATGGGCTTCTCTGGGGCCATGGCCCATACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||:::::   |||||||||||||||||||||||:::||||||:::|||||||||
Db        478 GGGAACCTCATGAATCTGACCTTCTGTGCTGACAACCTCATCAATCACTTCATGTGTGAT 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||||||||||||||||||||||||||   ::::::||||||   ||||||||||||
Db        538 ATTCTTCCCCTTCTTGAGCTCTCCTGCAATAGCACCTTCATAAATGAGCTCGTAATCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||         |||||||||||||||||||   :::||||||||||||
Db        598 ATTGTTGTGGCCTTTGATATTGGTGTGCCCATTGTCACCATCTTCATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              |||||||||||||||||||:::   |||   ||||||||||||||||||||||:::
Db        658 ATCCTCTCTAGCATTCTTCGCATGCATTCCACAGAGGGCAGGTCCAAGGCTTTCAGCACC 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::::|||   |||   |||||||||||||||||||||||   ||||||
Db        718 TGCAGCTCCCACTTGATTGTGGTTTGTCTTTTCTTTGGTTCTGGGGCTTTCATGTACCTC 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 AAACCTCCTTCCATTTTGCCTCTTGACCAAGGAAAAGTGTCTTCCTTGTTCTATATGATT 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||   ||||||||||||||||||||||||||||||||||||:::|||
Db        838 GTGGTGCCCATGTTGAACCCTCTGATCTATAGCTTGAGAAATAAGGATGTCAAAGTTGCT 897

Qy        301 LeuLysArgThrPheSerArg 307
              |||::::::|||      :::
Db        898 CTGAGGAAAACCTTGGGTAAG 918

RESULT 3
AY405611
```

```
LOCUS       AY405611                903 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Mus musculus OR8B8 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY405611
VERSION     AY405611.1  GI:39761585
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 903)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 903)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .903
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>903
                     /gene="OR8B8"
                     /locus_tag="HCM2281"
ORIGIN

Alignment Scores:
Pred. No.:              8.45e-89        Length:         903
Score:                  1158.00         Matches:        222
Percent Similarity:     84.00%          Conservative:   30
Best Local Similarity:  74.00%          Mismatches:     48
Query Match:            73.43%          Indels:         0
DB:                     29              Gaps:           0

US-10-023-597-24 (1-311) x AY405611 (1-903)

Qy         12 PheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGlnValProValPhePheLeuPhe 31
              |||||||||||||||||   ||||||||||:::::|||:::|||||||||||
Db          1 TTCATCCTTGCAGGCCTGACAGACCAGCCAGGACTCCGCATGCCCCTCTTCTTCCTGTTT 60

Qy         32 LeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIleLeuIleGlyLeu 51
              ||||||||||   ||||||||||||||||||||||||||||   ||||||||||
Db         61 CTAGGTTTCTACATGGTGACTGTGGTGGGGAACCTGGGTTTGATCTCCCTGATTGGGCTG 120

Qy         52 AsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSerLeuValAspPhe 71
              ||||||   ||||||   ||||||||||||||   ||||||||||||||:::|||||||
Db        121 AACTCTCACCTTCATACCCCTATGTACTTCTTTCTCTTCAATCTTTCCTTAATCGATTTC 180

Qy         72 SerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLysAsnIle 91
                 :::|||:::||||||   |||||||||||||||||||:::|||:::|||||||||
Db        181 TGCTATTCCTCTACTATCTCCCCAAAAATGCTGATGAGTTTTATCTCAAAGAAGAACATC 240

Qy         92 IleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPheSerGlu 111
              ||||||      |||||||||:::|||   ||||||||||||||||||||   ||||||
Db        241 ATCTCACATCCAGGGTGCATGGCACAGTTGTTTTTCTTCTGTTTTTTTGTTATCTCTGAG 300

Qy        112 SerPheIleLeuSerAlaMetValGluAspArgTyrValGlyIleCysAsnProLeuLeu 131
              ||||||||||||||||||||   ||||||||||||   ||||||||||||||:::
Db        301 TCTTTTATTCTGTCAGCCATGGCATATGACCGTTATGTTGCCATCTGTAACCCCCTGATG 360
```

```
Qy        132 TyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyrGlyMet 151
Db        361 TACATGGTCACCATGTCTCCCCAGGTTTGTTTACTCCTTTTGTTTGGTGTATATTTAATG 420

Qy        152 GlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAsp 171
Db        421 GGTTTTGTTGGAGCCATGGCCCACACCATATCCATGGCAAGGCTGACCTTCTGTGCTGAC 480

Qy        172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGly 191
Db        481 AACATTGTCAACCACTATATGTGTGATATCCTTCCCCTTCTGGAGCACTCCTGTACTAGC 540

Qy        192 SerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyValProIle 211
Db        541 ACCTATGTGAATGAACTGGTAGTCTTTATTTTTGTGAGTTTTGATATTGGTGTGCCCATT 600

Qy        212 ValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSerSerAla 231
Db        601 GTCACCATCTTCATTTCTTATGCCCTCATCCTCTCTAGCATTCTTCACATGCATTCCACA 660

Qy        232 GluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSerLeuPhe 251
Db        661 GAGGGCAGGTCCAAGGCCTTCAGCACCTGCAGCTCCCACATGATTGTGGTTTGTCTTTTC 720

Qy        252 PheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAspGlnGly 271
Db        721 TTTGGTTCTGGGGCTTTCATGTATCTCCAGCCACCTTCTGTTTTGTCCCTTGACCAAGGA 780

Qy        272 LysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIleTyrSer 291
Db        781 AAGGTGTCTTCCTTGTTCTATACGATTGTAGTACCTATGTTGAATCCTCTGATCTATAGC 840

Qy        292 LeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSerPheSer 311
Db        841 TTGAGAAATAAGGATGTCAAAGTTACTGTAAGGAAAACCTTGGACAGGAGAATATTTTCT 900

RESULT 4
AY405609
LOCUS       AY405609                903 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Homo sapiens OR8B8 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY405609
VERSION     AY405609.1  GI:39761583
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 903)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 903)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .903
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>903
                     /gene="OR8B8"
                     /locus_tag="HCM2281"
ORIGIN

Alignment Scores:
Pred. No.:              3.51e-87        Length:         903
Score:                  1139.00         Matches:        217
Percent Similarity:     83.67%          Conservative:   34
Best Local Similarity:  72.33%          Mismatches:     49
Query Match:            72.23%          Indels:         0
DB:                     29              Gaps:           0

US-10-023-597-24 (1-311) x AY405609 (1-903)

Qy         12 PheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGlnValProValPhePheLeuPhe 31
Db          1 TTTATCCTCGCAGGCTTAACTGACCAACCGGGAGTCCAGATCCCCCTCTTCTTCCTGTTT 60

Qy         32 LeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIleLeuIleGlyLeu 51
Db         61 CTAGGCTTCTACGTGGTCACTGTGGTGGGGAACCTGGGCTTGATAACCCTGATAAGGCTC 120

Qy         52 AsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSerLeuValAspPhe 71
Db        121 AACTCTCACTTGCACACCCCTATGTACTTCTTCCTCTATAACTTGTCCTTCATAGATTTC 180

Qy         72 SerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLysAsnIle 91
Db        181 TGCTATTCCAGTGTTATCACTCCCAAAATGCTGATGAGCTTTGTCTTAAAGAAGAACAGC 240

Qy         92 IleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPheSerGlu 111
Db        241 ATCTCCTACGCAGGGTGTATGACTCAGCTCTTCTTCTTTCTTTTCTTTGTTGTCTCTGAG 300

Qy        112 SerPheIleLeuSerAlaMetValGluAspArgTyrValGlyIleCysAsnProLeuLeu 131
Db        301 TCCTTCATCCTGTCAGCAATGGCGTACGACCGCTATGTGGCCATCTGTAACCCACTGTTG 360

Qy        132 TyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyrGlyMet 151
Db        361 TACATGGTCACCATGTCTCCCCAGGTGTGTTTTCTCCTTTTGTTGGGTGTCTATGGGATG 420

Qy        152 GlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAsp 171
Db        421 GGGTTTGCTGGGGCCATGGCCCACACAGCGTGCATGATGGGTGTGACCTTCTGTGCCAAT 480

Qy        172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGly 191
Db        481 AACCTTGTCAACCACTACATGTGTGACATCCTTCCCCTTCTTGAGTGTGCTTGCACCAGC 540

Qy        192 SerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyValProIle 211
Db        541 ACCTATGTGAATGAGCTTGTAGTGTTTGTTGTTGTGGGCATTGATATTGGTGTGCCCACA 600

Qy        212 ValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSerSerAla 231
Db        601 GTCACCATCTTCATTTCCTATGCTCTCATTCTCTCCAGCATCTTCCACATTGATTCCACG 660

Qy        232 GluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSerLeuPhe 251
Db        661 GAGGGCAGGTCCAAAGCCTTCAGCACCTGCAGCTCCCACATAATTGCAGTTTCTCTGTTC 720

Qy        252 PheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAspGlnGly 271
Db        721 TTTGGGTCAGGAGCATTCATGTACCTCAAACCCTTTTCTCTTTTAGCTATGAACCAGGGC 780

Qy        272 LysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIleTyrSer 291
Db        781 AAGGTGTCTTCCCTATTCTATACCACTGTGGTGCCCATGCTCAACCCATTAATTTATAGC 840

Qy        292 LeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSerPheSer 311
Db        841 CTGAGGAATAAGGACGTCAAAGTTGCTCTAAAGAAAATCTTGAACAAAAATGCATTCTCC 900
```

```
RESULT 5
AY407110
LOCUS       AY407110                799 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Pan troglodytes OR8B12 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY407110
VERSION     AY407110.1  GI:39763081
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 799)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .799
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>799
                     /gene="OR8B12"
                     /locus_tag="HCM2781"
ORIGIN

Alignment Scores:
Pred. No.:              9.3e-83       Length:         799
Score:                  1086.50       Matches:        213
Percent Similarity:     86.52%        Conservative:   18
Best Local Similarity:  79.78%        Mismatches:     35
Query Match:            68.90%        Indels:         1
DB:                     29            Gaps:           1

US-10-023-597-24 (1-311) x AY407110 (1-799)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||:::|||   |||||||||||||||||||||||||   ||||||   ||||||
Db          1 ATGGCAGCCAAAAAC---TCTTCTGTGACAGAGTTTATCCTCGAAGGCTTAACCCACCAG 57

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||::::::|||:::|||||||||||||||||||||||   ||||||||||||
Db         58 CCGGGACTGCGGATCCCCCTCTTCTTCCTGTTTCTGGGTTTCTACATGGTCACCGTGGTG 117

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||   ||||||||||||||      ||||||   |||||||||
Db        118 GGGAACCTGGGCTTGATAACCCTGATTGGGCTGAACTGTCACCTGCACACTCCCATGTAC 177

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||||||||   ||||||   ||||||:::||||||   ||||||
Db        178 TTCTTCCTTTTTAACCTCTCTTTAACAGATTTCTGTTTCTCCAGTACCATCACTCCCAAA 237

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        238 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              |||||||||||||||||||||||   |||||||||||:::|||||||||||||||||||
Db        298 CTCTTTTTCTTCTGCTTCTTTGTCATCTCTGAGTCCTTCGTCTTGTCAGCGATGGCGTAG 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   |||||||||||||||||||||||||:::|||||||||   ||||||
Db        358 GACCGCTATGTGGCCATCTGTAACCCACTGTTGTACACAGTCACCATGTCTTGCCAGGTG 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||   ||||||||||||       ||||||:::||||||||||
Db        418 TGTTTGCTCCTTTTGTTGGGTGCCTATGGGATGGGGTTTGCTGGGGCCATGGCCCACACA 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::|||:::   |||||||||||||||||||||||||||||||:::|||||||||
Db        478 GGAAGCATAATGAACCTGACCTTCTGTGCTGACAACCTTGTCAATCATTTCATGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||   ||||||:::|||   ||||||:::|||
Db        538 ATCCTTCCTCTCCTTGAGCTCTCCTGCAACAGCTCTTACATGAATGAGCTGGTGGTCTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||   |||   :::|||:::|||||||||   ||||||||||||||
Db        598 ATTGTGGTGGCTGTTGACGTTGGAATGCCCATTGTCACTGTCTTTATTTCTTATGCCCTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||      ||||||   ||||||||||||||||||||||:::
Db        658 ATCCTCTCCAGCATTCTACACAACAGTTCTACAGAAGGCAGGTCCAAAGCCTTTAGTACT 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::||||||   ||||||||||||||||||||||||||   ||||||
Db        718 TGCAGTTCCCACATAATTGTAGTTTCTCTTTTCTTTGGTTCTGGTGCTTTCATGTATCTC 777

Qy        261 LysProProSerIleLeuPro 267
              ||||||   |||||||||
Db        778 AAACCCCTTNCCATCCTGCCC 798

RESULT 6
AY405659
LOCUS       AY405659                930 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Mus musculus OR8B4 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY405659
VERSION     AY405659.1  GI:39761633
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>930
                     /gene="OR8B4"
```

/locus_tag="HCM2297"
ORIGIN

Alignment Scores:

Pred. No.:	2.02e-77	Length:	930
Score:	1024.50	Matches:	193
Percent Similarity:	77.74%	Conservative:	48
Best Local Similarity:	62.26%	Mismatches:	68
Query Match:	64.97%	Indels:	1
DB:	29	Gaps:	1

US-10-023-597-24 (1-311) x AY405659 (1-930)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||      ||||||||||||||||||||||||||||||    |||      :::|||
Db          1 ATGGCTCCAAGAAATAGCTCCTCTGTGACTGAATTTATCCTTGTGGGATTTTCAAACCAA 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   ||||||:::|||:::||||||:::||||||||||    |||   :::||||||:::
Db         61 CCAGCTCTCCAGCTGCCTCTCTTCTTCGTCTTCTTAGGTATCTATGTGCTGACTGTTATA 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||    ||||||||||||||||||    ||||||   ||||||||
Db        121 GGCAATTTGGGACTGATCACCTTAATTGGGTTGAATTCCAGTCTTCATACTCCAATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||   ||||||||||    :::||||||   :::|||                ||||||
Db        181 TTTTTCCTCTTCAACTTGTCCTTTATAGATTTCTGTTATTCCTGTGTGTTCACGCCAAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||      |||||||||   :::||||||||||||:::   ||||||||||:::|||
Db        241 ATGTTGAGTGATTTTGTCTCA---GAAAACATTATCTCCTACATGGGATGCATGACTCAA 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 ||||||||||||||||||    ||||||   ::::::|||   :::|||
Db        298 CTGTTTTTCTTCTGTTTCTTTGTCAATTCTGAATGCTACGTGCTGGTGTCAATGGCCTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||    ||||||||||||||||||||:::||||||||||||||||||||
Db        358 GATCGCTATGTGGCCATTTGCAATCCTCTGCTGTACACAGTTACCATGTCTCCTCAGGTC 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||   ||||||:::       |||   :::|||      ||||||:::|||||||||
Db        418 TGCACTCTGCTGATGTTTTGTTCATATGTGATAGGGTTTGCTGGAGCCATGGCCCATACT 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||:::::::::    ||||||||||||      |||:::::::::||||||:::|||:::
Db        478 GGAAGCATGCTGACACTTACTTTCTGTGATTCCAATATGATCCACCATTATCTCTGTGAA 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              :::|||||||||||||:::||||||||      :::|||   |||   ||||||   |||
Db        538 GTTCTTCCCCTTTTACAGCTCTCCTGCACCAGTACCTATGCCAACGAGCTTGTGTTTTTT 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||    |||   |||              :::::::::||||||||||
Db        598 ATTGTTGTTGGAGTAGTCATCACAGCATCCAGCATCAGCATCTTTATCTCTTATGCATTG 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              |||||||||:::||||||::::::   ||||||||||||||||||||||||||   :::
Db        658 ATTCTCTCCAACATTCTTAAGATTCCTTCTGCTGAAGGTAGATCCAAAGCCTTTGGCACA 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||   |||:::::::::||||||:::||||||||||||||||||||||||||||||||
Db        718 TGTGGTTCCCATGTAGTTGCTGTTGCTCTGTTTTTTGGGTCAGGGGCATTCACCTATTTA 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
                          :::::::::|||:::   :::|||:::||||||||||
Db        778 ACAACCTCTTTCCCTGGTTCAATGGAAGAGGGAAGGTTTGCTTCAGTATTTTATACCAAT 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        838 GTGGTTCCCATGCTTAACCCACTGATCTATAGTTTGAGGAATAAAGACGTTAAACTTGCC 897
```

```
Qy        301 LeuLysArgThrPheSerArgIleSerPhe 310
              |||   :::|||      |||:::   |||
Db        898 CTAAATAAAACCTTGAAGAGAGTGCTCTTC 927
```

RESULT 7
AY405657

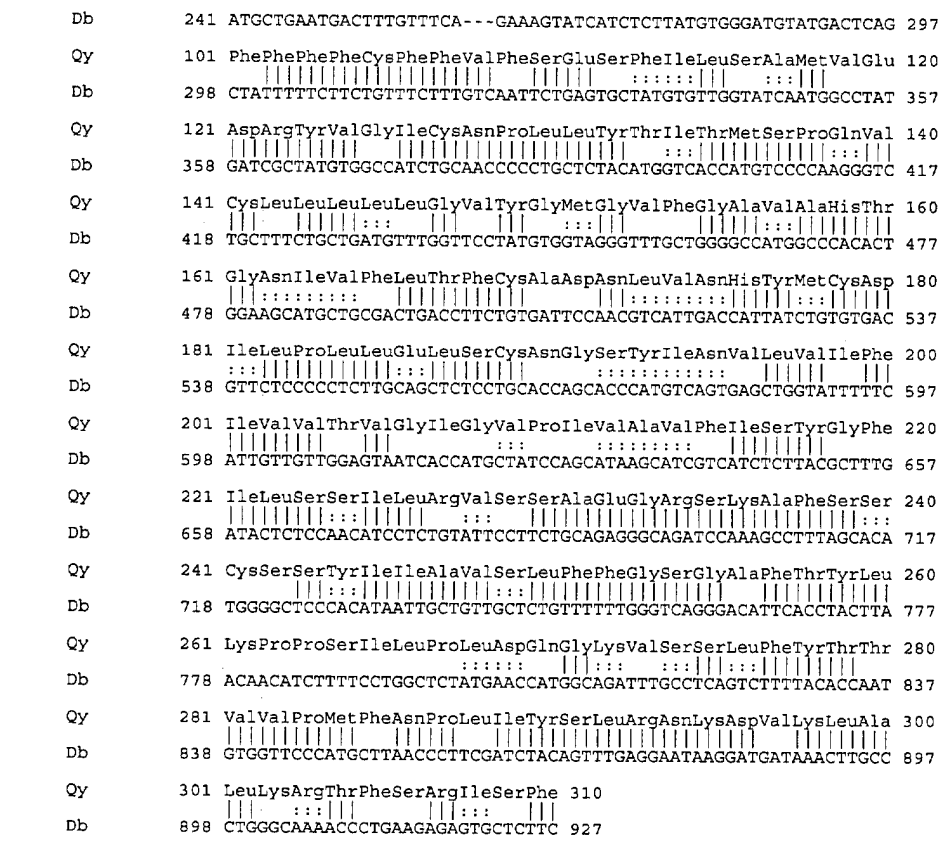
```
LOCUS       AY405657                930 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Homo sapiens OR8B4 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY405657
VERSION     AY405657.1  GI:39761631
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>930
                     /gene="OR8B4"
                     /locus_tag="HCM2297"
ORIGIN
```

Alignment Scores:

Pred. No.:	8.42e-72	Length:	930
Score:	958.50	Matches:	186
Percent Similarity:	74.84%	Conservative:	46
Best Local Similarity:	60.00%	Mismatches:	77
Query Match:	60.78%	Indels:	1
DB:	29	Gaps:	1

US-10-023-597-24 (1-311) x AY405657 (1-930)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||         ||||||||||||||||||||||||||||||    ||||||      |||
Db          1 ATGACTCTGAGAAACAGCTCCTCAGTGACTGAGTTTATCCTTGTGGGATTATCAGAACAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||:::|||:::|||    ||||||||||    |||      ||||||||||
Db         61 CCAGGGCTCCAGCTCCCTCTTTTCCTTCTATTCTTAGGGATCTATGTGTTCACTGTGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||    |||||||||:::|||      ||||||   ||||||||||
Db        121 GGCAACTTGGGCTTGATCACCTTAATTGGGATAAATCCTAGCCTTCACACCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||    ||||||||||||   :::|||      :::|||                ||||||
Db        181 TTTTTCCTCTTCAACTTGTCCTTTATAGATCTCTGTTATTCCTGTGTGTTTACCCCCAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||      |||||||||   ::::::|||||||||:::   ||||||||||:::|||
```



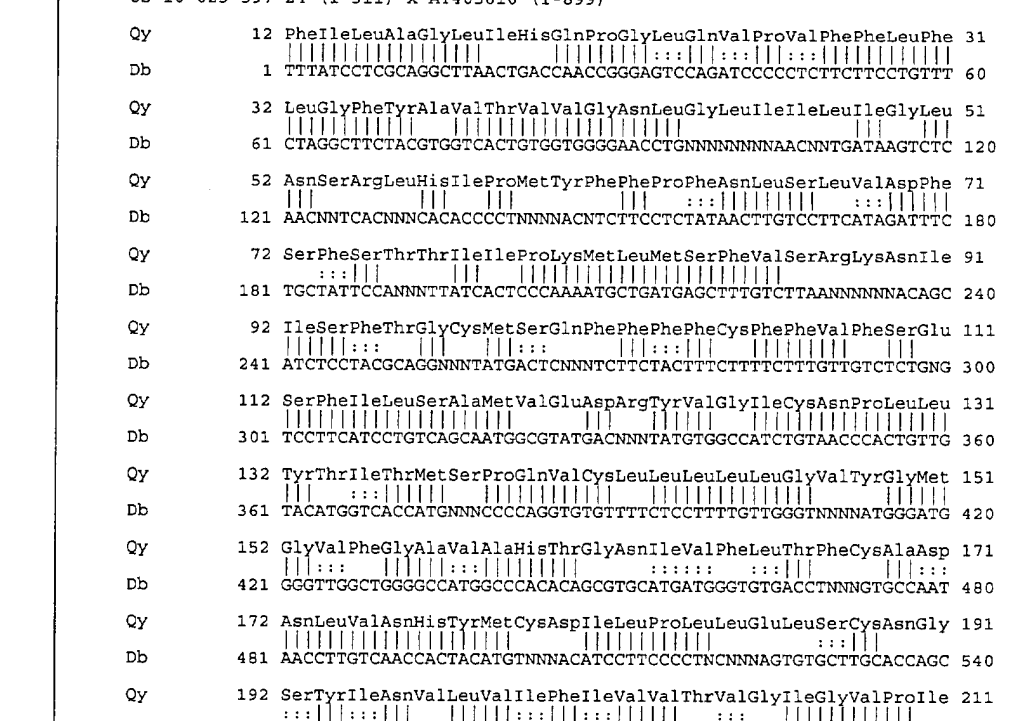
```
Db        241 ATGCTGAATGACTTTGTTTCA---GAAAGTATCATCTCTTATGTGGGATGTATGACTCAG 297

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        298 CTATTTTTCTTCTGTTTCTTTGTCAATTCTGAGTGCTATGTGTTGGTATCAATGGCCTAT 357

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        358 GATCGCTATGTGGCCATCTGCAACCCCCTGCTCTACATGGTCACCATGTCCCCAAGGGTC 417

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        418 TGCTTTCTGCTGATGTTTGGTTCCTATGTGGTAGGGTTTGCTGGGGCCATGGCCCACACT 477

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        478 GGAAGCATGCTGCGACTGACCTTCTGTGATTCCAACGTCATTGACCATTATCTGTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        538 GTTCTCCCCCTCTTGCAGCTCTCCTGCACCAGCACCCATGTCAGTGAGCTGGTATTTTTC 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        598 ATTGTTGTTGGAGTAATCACCATGCTATCCAGCATAAGCATCGTCATCTCTTACGCTTTG 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        658 ATACTCTCCAACATCCTCTGTATTCCTTCTGCAGAGGGCAGATCCAAAGCCTTTAGCACA 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        718 TGGGGCTCCCACATAATTGCTGTTGCTCTGTTTTTTGGGTCAGGGACATTCACCTACTTA 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        778 ACAACATCTTTTCCTGGCTCTATGAACCATGGCAGATTTGCCTCAGTCTTTTACACCAAT 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        838 GTGGTTCCCATGCTTAACCCTTCGATCTACAGTTTGAGGAATAAGGATGATAAACTTGCC 897

Qy        301 LeuLysArgThrPheSerArgIleSerPhe 310

Db        898 CTGGGCAAAACCCTGAAGAGAGTGCTCTTC 927
```

RESULT 8
AY405610
LOCUS AY405610 899 bp DNA linear GSS 16-DEC-2003
DEFINITION Pan troglodytes OR8B8 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY405610
VERSION AY405610.1 GI:39761584
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 899)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 899)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .899
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1. .>899
/gene="OR8B8"
/locus_tag="HCM2281"
ORIGIN

Alignment Scores:

Pred. No.:	8.19e-70	Length:	899
Score:	935.00	Matches:	184
Percent Similarity:	72.91%	Conservative:	34
Best Local Similarity:	61.54%	Mismatches:	81
Query Match:	59.29%	Indels:	0
DB:	29	Gaps:	0

US-10-023-597-24 (1-311) x AY405610 (1-899)



```
Qy         12 PheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGlnValProValPhePheLeuPhe 31

Db          1 TTTATCCTCGCAGGCTTAACTGACCAACCGGGAGTCCAGATCCCCCTCTTCTTCCTGTTT 60

Qy         32 LeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIleLeuIleGlyLeu 51

Db         61 CTAGGCTTCTACGTGGTCACTGTGGTGGGGAACCTGNNNNNNNNAACNNTGATAAGTCTC 120

Qy         52 AsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSerLeuValAspPhe 71

Db        121 AACNNTCACNNNCACACCCCTNNNNACNTCTTCCTCTATAACTTGTCCTTCATAGATTTC 180

Qy         72 SerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLysAsnIle 91

Db        181 TGCTATTCCANNNTTATCACTCCCAAAATGCTGATGAGCTTTGTCTTAANNNNNNACAGC 240

Qy         92 IleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPheSerGlu 111

Db        241 ATCTCCTACGCAGGNNNTATGACTCNNNTCTTCTACTTTCTTTTCTTTGTTGTCTCTGNG 300

Qy        112 SerPheIleLeuSerAlaMetValGluAspArgTyrValGlyIleCysAsnProLeuLeu 131

Db        301 TCCTTCATCCTGTCAGCAATGGCGTATGACNNNTATGTGGCCATCTGTAACCCACTGTTG 360

Qy        132 TyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyrGlyMet 151

Db        361 TACATGGTCACCATGNNNCCCCAGGTGTGTTTTCTCCTTTTGTTGGGTNNNNATGGGATG 420

Qy        152 GlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAsp 171

Db        421 GGGTTGGCTGGGGCCATGGCCCACACAGCGTGCATGATGGGTGTGACCTNNNGTGCCAAT 480

Qy        172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGly 191

Db        481 AACCTTGTCAACCACTACATGTNNNACATCCTTCCCCTNCNNNAGTGTGCTTGCACCAGC 540

Qy        192 SerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyValProIle 211

Db        541 ACCTATGTGAATGAGCTTGTAGTGTTTGTTGTTGTGGGCATTGATATTGGTGTGCCCACA 600

Qy        212 ValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSerSerAla 231

Db        601 GTCACCATCTTCATTTCCTATGCTCTCATTCTNNNCAGCATCTTCCACATTGATTCCACG 660

Qy        232 GluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSerLeuPhe 251

Db        661 GAGGGCAGNTCCAAAGCCTTCAGCACCTGCAGCTCCCACATAATTGCAGTTTCTCTNNNC 720

Qy        252 PheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAspGlnGly 271
```

```
Db        721 TTTGGGTCAGGAGCATTCATGTATCTCAAACCCTTTTCTCTTTTAGCTATGAACCAGGGC 780

Qy        272 LysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIleTyrSer 291

Db        781 AAGNNNTCTTTCCTATTCTATACCACTGTGGTGCCCATGCTCAACCCATTAATTTATAGC 840

Qy        292 LeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSerPhe 310

Db        841 CTGNGGAATAAGGACGTCAAAGTTNCNCTAAAGAAANNNNNNAACAAAAATGCATTC 897

RESULT 9
AY419057
LOCUS       AY419057                930 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Mus musculus OR8A1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY419057
VERSION     AY419057.1  GI:39775017
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>930
                     /gene="OR8A1"
                     /locus_tag="HCM6763"
ORIGIN

Alignment Scores:
Pred. No.:              8.04e-69        Length:         930
Score:                  923.50          Matches:        180
Percent Similarity:     74.19%          Conservative:   50
Best Local Similarity:  58.06%          Mismatches:     79
Query Match:            58.56%          Indels:         1
DB:                     29              Gaps:           1

US-10-023-597-24 (1-311) x AY419057 (1-930)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db          1 ATGGCTGCAGAGAATCAATCTACAGTGACAGAGTTTATCCTCAGAGGGCTAACTAACAGG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

Db         61 CCAGAGCTCCAGCTACCCCTCTTGCTCCTCTTCCTGGGGATCTACATAATCACCATGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db        121 GGAAATCTGGGCATGATCACCTTGATTGGACTCAACTCACAGCTTCACACCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80

Db        181 TTCTTCCTCAGTAACCTGTCACTCGTGGATCTCTGCTACTCCTCTGTCATTACCCCTAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db        241 ATGCTCATCAACTTTGTGTCCCAAAGAAACCTCATCTCCTATGTGGGGTGCATGTCTCAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

Db        301 CTATACTTCTTCCTGGTTTTCGTCATTGCTGAGTGTTACATGCTCACTGTGATGGCCTAT 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db        361 GACCGCTATGTGGCCATCTGTCAACCCTTGCTTTATAACATCATCATGTCTCCTGCCCTC 420

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db        421 TGCTCTCTGCTGGTGGTTTTTGTCTATGCCATGGGACTCATAGGCTCAACAATAGAGACT 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db        481 AGCCTCATGTTGAAATTAAACTATTGTGAGGAC---CTCATCAGCCACTACTTCTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db        538 ATCCTCCCCCTCATGAAGCTCTCTTGCTCTAGTACCTATGATATAGAAATGGCTGTCTTC 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

Db        598 TTTTTAGCTGGGTTCAACATTATAGTTACCAGCTTAACCGTACTCATTTCCTATGCATTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240

Db        658 ATCCTGTCCAGCATCCTTCGCATCAGCTCCAATGAGGGCAGGTCCAAAGCCTTCAGCACC 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db        718 TGCAGCTCCCACTTTGCAGCTGTGGGCTTGTTCTATGGATCCACAGCATTCATGTACTTA 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

Db        778 AAGCCTTCTACAGCTAGTTCCCTGGCCCAGGAGAATGTAGCTTCTGTGTTCTACACCACA 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

Db        838 GTGATCCCCATGCTCAACCCCCTGATCTACAGCCTGAGAAACAAGGAGGTGAAGACCGCA 897

Qy        301 LeuLysArgThrPheSerArgIleSerPhe 310

Db        898 CTGGACAAAACACTGAGGAGAAGAGTCTTT 927

RESULT 10
AY419055
LOCUS       AY419055                930 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Homo sapiens OR8A1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY419055
VERSION     AY419055.1  GI:39775015
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
```

```
   PUBMED   14671302
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>930
                     /gene="OR8A1"
                     /locus_tag="HCM6763"
ORIGIN

Alignment Scores:
Pred. No.:              7.68e-66        Length:         930
Score:                  888.50          Matches:        172
Percent Similarity:     74.67%          Conservative:   55
Best Local Similarity:  56.58%          Mismatches:     76
Query Match:            56.34%          Indels:         1
DB:                     29              Gaps:           1

US-10-023-597-24 (1-311) x AY419055 (1-930)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||   |||   |||:::|||||||||||||||||||   ||||||      :::
Db          1 ATGGCTGCAGGAAATCACTCTACAGTGACAGAGTTCATTCTCAAGGGTTTAACGAAGAGA 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
                  ||||||:::|||:::|||   ||||||||||   |||   ||||||:::|||
Db         61 GCAGACCTCCAGCTCCCCCTCTTTCTCCTCTTCCTCGGGATCTACTTGGTCACCATCGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||:::|||   ||||||   |||||||||:::||||||   |||||||||
Db        121 GGGAACCTGGGCATGATCACTCTAATTTGTCTGAACTCTCAGCTGCACACCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              :::|||      ||||||||||||:::|||      :::|||:::   |||   ||||||
Db        181 TACTTTCTCAGCAATCTGTCACTCATGGATCTCTGCTACTCCTCCGTCATTACCCCTAAG 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||:::::|||||||||   ||||||||||||||:::   ||||||||||||||||
Db        241 ATGCTGGTGAACTTTGTGTCAGAGAAAAACATCATCTCCTACGCAGGGTGCATGTCACAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 :::||||||      ||||||   :::|||   :::::|||:::   |||
Db        301 CTCTACTTCTTCCTTGTTTTTGTCATTGCTGAGTGTTACATGCTGACAGTGATGGCCTAC 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||:::||||||||||||   |||   ||||||
Db        361 GACCGCTATGTTGCCATCTGCCACCCTTTGCTTTACAACATCATTATGTCTCATCACACC 420

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||:::      ||||||   :::|||:::   |||:::         |||
Db        421 TGCCTGCTGCTGGTGGCTGTGGTCTACGCCATCGGACTCATTGGCTCCACAATAGAAACT 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||   ::::::   |||   :::|||   ::::::|||:::::||||||   ||||||
Db        481 GGCCTCATGTTAAAACTGCCCTATTGT---GAGCACCTCATCAGTCACTACTTCTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||:::::|||||||||:::   :::|||         :::   :::|||
Db        538 ATCCTCCCTCTCATGAAGCTGTCCTGCTCTAGCACCTATGATGTTGAGATGACAGTCTTC 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
                 :::            |||   |||      :::   |||   :::||||||   |||

Db        598 TTTTTGGCTGGATTCAACATCATAGTCACGAGCTTAACAGTTCTTGTTTCTTACACCTTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||   :::|||:::   ||||||||||||||||||||||||:::
Db        658 ATTCTCTCCAGCATCCTCGGCATCAGCACCACAGAGGGGAGATCCAAAGCCTTCAGCACC 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::::   ||||||   :::|||:::||||||   ||||||   ||||||
Db        718 TGCAGCTCCCACCTTGCAGCCGTGGGAATGTTCTATGGATCAACTGCATTCATGTACTTA 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   :::|||      |||   |||      |||:::|||:::||||||||||||
Db        778 AAACCCTCCACAATCAGTTCCTTGACCCAGGAGAATGTGGCCTCTGTGTTCTACACCACG 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              |||:::||||||   ||||||||||||||||||||||||||||:::||||||   |||
Db        838 GTAATCCCCATGTTGAATCCCCTAATCTACAGCCTGAGGAACAAGGAAGTAAAGGCTGCC 897

Qy        301 LeuLysArgThr 304
              :::::::::|||
Db        898 GTGCAGAAAACG 909

RESULT 11
AY419056
LOCUS       AY419056                930 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Pan troglodytes OR8A1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY419056
VERSION     AY419056.1  GI:39775016
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>930
                     /gene="OR8A1"
                     /locus_tag="HCM6763"
ORIGIN

Alignment Scores:
Pred. No.:              3.69e-65        Length:         930
Score:                  880.50          Matches:        171
Percent Similarity:     74.34%          Conservative:   55
Best Local Similarity:  56.25%          Mismatches:     77
Query Match:            55.83%          Indels:         1
DB:                     29              Gaps:           1

US-10-023-597-24 (1-311) x AY419056 (1-930)
```

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||   |||   |||:::||||||||||||||||||   ||||||      :::
Db          1 ATGGCTGCAGGAAATCACTCTACAGTGACAGAATTCATTCTCAAGGGTTTAANNNAGAGA 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
                  ||||||:::|||:::|||   ||||||||||||   |||   |||||:::::::
Db         61 GCAGACCTCCAGCTCCCCCTCTTTCTCCTCTTCCTCGGGATCTACTTGGTCACCATCATG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||:::|||   ||||||   |||||||||:::||||||   |||||||||
Db        121 GGGAACCTGGGCATGATCACTCTAATTTGTCTGAACTCTCAGCTGCACACCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              :::|||      ||||||||||||||||||      :::|||:::   |||   ||||||
Db        181 TACTTCCTCAGCAATCTGTCACTCGTGGATCTCTGCTACTCCTCCGTCATTACCCCTAAG 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||:::::|||||||||   ||||||||||||||:::   ||||||||||||||||
Db        241 ATGCTGGTGAACTTTGTGTCAGAGAAAAACATCATCTCCTACGCAGGGTGCATGTCACAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
                 :::||||||      ||||||   :::|||   ::::::|||:::   |||
Db        301 CTCTACTTCTTCCTTGTTTTTGTCATTGCTGAGTGTTACATGCTGACAGTGATGGCCTAC 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||   ||||||:::||||||||||||   |||   ||||||
Db        361 GACCGCTATGTTGCCATCTGCCACCCTTTGCTTTACAACATCATTATGTCTCATCACACC 420

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||   ||||||:::      ||||||   :::|||:::   |||:::          |||
Db        421 TGCCCGCTGCTGGTGGCTGTGGTCTACGCCATCGGACTCATTGGCTCCACAATAGAAACT 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||   ::::::   |||   :::|||   ::::::|||::::::||||||   ||||||
Db        481 GGCCTCATGTTAAAACTGCCCTATTGT---GAGCACCTTATCAGTCACTACTTCTGTGAC 537

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||:::::||||||||:::   :::|||          :::   :::|||
Db        538 ATCCTCCCTCTCATGAAGCTGTCCTGCTCTAGCACCTATGATGTTGAGATGACAGTCTTC 597

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
                 :::          |||   |||      :::   |||   :::||||||   |||
Db        598 TTTTTGGCTGGATTCGACATCATAGTCACGAGCTTAACAGTTCTTGTTTCTTACACCTTC 657

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||   :::|||:::   ||||||||||||||||||||||||:::
Db        658 ATTCTCTCCAGCATCCTCGGCATCAGCACCACAGAGGGGAGATCCAAAGCCTTCAGCACC 717

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||:::::   ||||||   :::|||:::||||||   ||||||   ||||||
Db        718 TGCAGCTCCCACCTTGCAGCCGTGGGAATGTTCTATGGATCAACTGCATTCATGTACTTA 777

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||   :::|||      |||   |||      |||:::|||:::||||||||||||
Db        778 AAACCCTCCACAATCAGTTCCTTGACCCAGGAGAATGTGGCCTCTGTGTTCTACACCACG 837

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              |||:::||||||   ||||||||||||||||||||||||||||:::||||||   |||
Db        838 GTAATCCCCATGTTGAATCCCCTAATCTACAGCCTGAGGAACAAGGAAGTAAAGGCTGCC 897

Qy        301 LeuLysArgThr 304
              ::::::::::|||
Db        898 GTGCAGAAAACG 909

RESULT 12
AZ255734
LOCUS       AZ255734                726 bp    DNA     linear   GSS 26-JUL-2000
DEFINITION  RPCI-23-166I9.TJ RPCI-23 Mus musculus genomic clone RPCI-23-166I9,
            genomic survey sequence.
ACCESSION   AZ255734
VERSION     AZ255734.1  GI:9458784
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-166I9.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 166  row: I  column: 9
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .726
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="RPCI-23-166I9"
                      /sex="Female"
                      /lab_host="DH10B"
                      /clone_lib="RPCI-23"
                      /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                      EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                      brain genomic DNA was isolated and  partially digested
                      with a combination of EcoRI and EcoRI Methylase.   Size
                      selected DNA was cloned into the pBACe3.6 vector at the
                      EcoRI  sites.  The ligation products were transformed into
                      DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

Alignment Scores:
Pred. No.:                2.02e-59        Length:          726
Score:                    812.00          Matches:         149
Percent Similarity:       78.01%          Conservative:    39
Best Local Similarity:    61.83%          Mismatches:      53
Query Match:              51.49%          Indels:          0
DB:                       28              Gaps:            0

US-10-023-597-24 (1-311) x AZ255734 (1-726)

Qy         25 ValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGly 44
              :::|||:::|||||||||||||||||      |||   ||||||:::   ||||||||||
Db          2 CTCCCTCTGTTCTTCTTGTTCCTTGTAATGTATATAGTAACTATGACCGGGAATTTGGGT 61

Qy         45 LeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPhe 64
              |||:::|||||||||||||||||||||   ||||||   ||||||||||||||   |||
Db         62 TTGGTAATTCTAATTGGACTGAATTCTCACCTTCACACCCCGATGTACTTTTTTCTCTTT 121

Qy         65 AsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMetSer 84
              ||||||||||:::|||      :::|||:::          ||||||||||:::::
Db        122 AACTTGTCTTTGATAGACCTCTGTTACTCTTCAGTGTTTACACCCAAAATGCTGTTGAAC 181

Qy         85 PheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePhe 104
              |||:::      ||||||||||||||:::||||||||||||:::|||   :::|||:::
Db        182 TTCATATTAAATAAGAATATTATCTCTTATACGGGGTGTATGACCCAACTCTATTTTTAT 241

Qy        105 CysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyrVal 124
                  |||||||||   ||||||   :::::|||   :::|||      ||||||||||||
```

```
Db        242 TCATTCTTTGTTATTTCTGAGTGTTATGTGTTAATGTCAATGGCCTATGATCGCTATGTA 301

Qy        125 GlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeu 144
              ||||||||||||||||||||    |||   |||:::|||:::::|||     |||
Db        302 GCCATCTGTAATCCACTCTTATATAATATTGCTATGACACCTAAAATATGTTCCTATCTT 361

Qy        145 LeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleVal 164
              :::|||||   |||   |||       ||||||:::||||||||||||   ::::::
Db        362 ATGCTTGGTTCATACTTGATGGCATTTTCTGGTGCCATGGCTCACACAGGATGTATGTTA 421

Qy        165 PheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeu 184
              ||||||||||||      |||   :::|||||||||   ||||||||||||||:::
Db        422 AGATTGACCTTCTGTGATGCAAACACCATCAATCACTACTTCTGTGACATCCTCCCTGTG 481

Qy        185 LeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePheIleValValThr 204
              ::::::|||||||||      :::|||:::|||   |||   :::|||||||||||
Db        482 ATGCAGCTCTCCTGCACCAGCACCTATGTCAATGAACTGGAAGTTTTCATTGTCGTGGGC 541

Qy        205 ValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSer 224
              :::   |||   ||||||   :::   :::|||||||||||||||||||||||||||||
Db        542 ATCAATATCCTTGTGCCCAGCATCACCATTTTTATCTCTTATGGGTTCATTCTCTCCAGC 601

Qy        225 IleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyr 244
              |||      :::::::|||   |||||||||||||||||||||:::|||||||||:::
Db        602 ATTTTCCACATCAACTCCAATGAGGGCAGGTCCAAGGCCTTCAGCACCTGCAGTTCCCAC 661

Qy        245 IleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSer 264
              |||||||||   ||||||||||||   ||||||||||||   ||||||||||||   |||
Db        662 ATAATTGCTGGTTCTCTCTTCTTTGCATCAGGTGCATTTATGTATCTTAAACCCTCCTCA 721

Qy        265 Ile 265
              :::
Db        722 GTT 724
```

RESULT 13
CB174479
LOCUS CB174479 748 bp mRNA linear EST 09-OCT-2003
DEFINITION OR_2oE1H11_020911.y1 Embryonic mouse olfactory epithelium library Mus musculus cDNA clone 2oE1H11 5', mRNA sequence.
ACCESSION CB174479
VERSION CB174479.1 GI:37593108
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 748)
AUTHORS Young,J.M., Shykind,B.M., Lane,R.P., Tonnes-Priddy,L., Ross,J.A., Walker,M., Williams,E.M. and Trask,B.J.
TITLE Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels
JOURNAL Genome Biol. 4 (11), R71.1-R71.15 (2003)
COMMENT Contact: Young JM
Trask Lab, Division of Human Biology
Fred Hutchinson Cancer Research Center
1100 Fairview Avenue N., C3-168, P.O. Box 19024, Seattle, WA 98109-1024, USA
Tel: 206 667 1471
Fax: 206 667 6524
Email: jayoung@fhcrc.org
Young gene new name GA_x6K02T2PVTD-32156773-32157705
Young gene old name GA_x5J8B7W60AJ-1146353-1145421
Zhang gene name MOR167-3
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .748
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Mixed Swiss-Webster and C57BL/6"
/db_xref="taxon:10090"
/clone="2oE1H11"
/tissue_type="Olfactory and respiratory epithelium"
/dev_stage="E16.5-E18.5"
/clone_lib="Embryonic mouse olfactory epithelium library"
/note="Organ: Olfactory turbinates; Vector: LambdaZAPII-XR; Site_1: EcoRI; Site_2: XhoI; This library was provided by Tyler Cutforth. mRNA was prepared from the olfactory and respiratory epithelium of mixed Swiss-Webster and C57BL/6 embryos, aged E16.5-E18.5. Oligo-dT primed cDNA was directionally cloned into Stratagene's lambdaZAPII-XR vector."
ORIGIN

Alignment Scores:

Pred. No.:	1.27e-57	Length:	748
Score:	791.00	Matches:	149
Percent Similarity:	78.79%	Conservative:	33
Best Local Similarity:	64.50%	Mismatches:	49
Query Match:	50.16%	Indels:	1
DB:	14	Gaps:	0

US-10-023-597-24 (1-311) x CB174479 (1-748)

```
Qy         80 LysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSer 99
              ||||||||||||:::||:::      ||||||   ||||||:::   ||||||::::::
Db          3 AAAATGCTGATGAACTTTATACTGAGCAAGAATGCAATTTCTTACATGGGATGTTTAACC 62

Qy        100 GlnPhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetVal 119
              |||   :::||||||||||||||||||   ||||||   ::::::|||::::::|||
Db         63 CAGCTATATTTCTTCTGCTTTTTTGTCATTTCTGAGTGTTATGTGTTGACTTCAATGGCT 122

Qy        120 GluAspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGln 139
              ||||||||||||   |||||||||||||||||||||||:::   ||||||||::::
Db        123 TATGATCGCTATGTGGCCATCTGTAATCCACTCTTGTACACGGTTGCCATGTCCCCTAAA 182

Qy        140 ValCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHis 159
              :::||||||   |||:::||||||   |||   |||   ||||||:::||||||
Db        183 TTGTGTTTGAACCTTATGCTTGGTACATATGCAATGGCATTTTCTGGTGCCATGGCTCAC 242

Qy        160 ThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCys 179
              ||||||      :::   ||||||||||||      |||   :::|||||||   |||
Db        243 ACAGGATGCAGNCTGAGACTGACCTTCTGTGATGCTAACACCATCAACCACTACTTCTGT 302

Qy        180 AspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIle 199
              ||||||||||||::::::::|||||||||      :::|||:::|||   ||||||:::
Db        303 GACATCCTCCCTGTGATGCAGCTCTCCTGCACCAGCACCTATGTCAATGAGCTTGTAGTT 362

Qy        200 PheIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGly 219
              ||||||||||||   :::   |||   ||||||   :::   :::||||||||||||||
Db        363 TTCATTGTTGTGGGCATCAATATCATTGTGCCAAGCATCACTATCTTCATCTCTTATGGC 422

Qy        220 PheIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSer 239
              ||||||||||||||||||      :::   |||   ||||||||||||||||||||||
Db        423 TTCATCCTCTCCAGCATTTTTCACATCAAGTCCAATGAAGGCAGGTCCAAGGCCTTCAGC 482

Qy        240 SerCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyr 259
              :::||||||||:::||||||||||||   ||||||||||||||||||||||   |||
Db        483 ACCTGCAGTTCCCACATAATTGCAGTTTGTCTCTTCTTTGGATCAGGTGCATTTATGTAT 542

Qy        260 LeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThr 279
              |||||||||   |||          :::||||||||||||   ||||||:::||||||||
Db        543 CTCAAACCATCTTCTTCTTCATCTATGGATCAAGGGAAAACCTCTTCTGTGTTTTATACA 602

Qy        280 ThrValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeu 299
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||:::
Db        603 AATGTGGTTCCCATGATGAATCCCTTAATCTACAGTTTAAGGAACAAAGATGTCAAGATA 662

Qy        300 AlaLeuLysArgThrPheSerArgIleSerPhe 310
              ||||||   :::|||   ||||||      |||
Db        663 GCCCTGAG-AAAACCCTGAGCAGATGGAAATTT 694
```

RESULT 14
AY405614
LOCUS AY405614 933 bp DNA linear GSS 16-DEC-2003
DEFINITION Mus musculus OR5AR1 gene, VIRTUAL TRANSCRIPT, partial sequence,

```
            genomic survey sequence.
ACCESSION   AY405614
VERSION     AY405614.1  GI:39761588
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>933
                     /gene="OR5AR1"
                     /locus_tag="HCM2282"
ORIGIN

Alignment Scores:
Pred. No.:              4.12e-57        Length:         933
Score:                  786.00          Matches:        151
Percent Similarity:     68.40%          Conservative:   59
Best Local Similarity:  49.19%          Mismatches:     97
Query Match:            49.84%          Indels:         0
DB:                     29              Gaps:           0

US-10-023-597-24 (1-311) x AY405614 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||      ||||||   |||   ||||||||||||:::      |||:::
Db          1 ATGGATAAAGAAAATCACTCAGTTGTGACTGAATTTGTCTTTATGGGCATAACTCAAGAC 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   ||||||:::       |||   :::||||||      |||   |||   |||:::
Db         61 CCTCAGCTGCAGATCATCTTTTTTGTGGTCTTCCTCCTTGTCTACCTGGTCAATGTGATA 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||:::|||:::||||||||||||      :::|||:::||||||   |||||||||
Db        121 GGGAATGTTGGTATGATCATCCTGATCATAACAGACAGTCAGCTTCACACCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||      |||||||||   ||||||      :::|||:::   |||   |||:::
Db        181 TTTTTCCTCTGCAACCTCTCCTTTGTTGACCTGGGCTACTCATCAGCTATTGCTCCCAGG 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||      |||:::::::::      :::||||||||||:::   |||   :::|||
Db        241 ATGCTGGCTGATTTTCTAACAAAGCACAAAGTTATCTCTTTTTCCAGCTGTGCCACCCAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              |||   ||||||      ||||||   :::|||   :::::::|||:::||||||
Db        301 TTTGCTTTCTTTGTAGGCTTTGTGGATGCTGAGTGCTATGTCCTGGCTGCAATGGCCTAT 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||:::|||   ||||||   ||||||   |||:::      ||||||   :::|||

Db        361 GACCGCTTTGTGGCCATCTGCAGACCTCTTCACTATAGCACTCTCATGTCCAAGAAAGTC 420

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||:::|||:::||||||   |||      |||:::          ||||||||||||
Db        421 TGCTTGGTCCTAATGCTGGGCTCTTACTTCGCTGGCCTAGTGAGTTTAGTGGCGCATACT 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
                          |||:::::::|||      |||::::::|||||:::   |||:::
Db        481 TCTCTCACTTTCAGCCTCAGTTACTGTGGTTCCAATATCATCAACCATTTCTTCTGTGAA 540

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||   |||||||||   |||||||||:::   :::||||||:::   :::::::::|||
Db        541 ATCCCACCACTCTTAGCTCTCTCTTGCTCAGATACCTACATTAGTGAGATACTGCTGTTT 600

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
                 :::                  |||:::   :::||||||||||||   |||
Db        601 AGTTTGTGTGGCTTCATTGAGTTCAGCACCATCCTTATCATCTTCATCTCCTATGCCTTC 660

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||   :::|||:::|||:::   ||||||||||||||   ||||||||||:::
Db        661 ATTCTCATTGCCATCATCAGAATTCGATCAGCTGAAGGCCGGCTAAAGGCTTTTTCCACC 720

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||   |||::::::      |||:::||||||:::|||:::      |||   ||||||
Db        721 TGCGGGTCTCATCTTACTGGTGTTACCCTTTTTTATGGTACAGTCATGTTCATGTATCTG 780

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              :::|||   |||          |||||||||   |||   :::|||:::||||||||||
Db        781 AGGCCAACATCCAGTTACTCCCTGGACCAAGACAAGTGGGCCTCTGTGTTCTACACAATT 840

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ::::::||||||   ||||||||||||||||||||||||||||||||||||||||   |||
Db        841 ATCATTCCTATGCTGAACCCCTTGATCTACAGTTTGCGGAACAAGGATGTGAAAGCCGCT 900

Qy        301 LeuLysArgThrPheSerArg 307
                 |||:::         :::
Db        901 TTCAAGAAACTGATTGGAAAG 921

RESULT 15
AY405612
LOCUS       AY405612                933 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Homo sapiens OR5AR1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY405612
VERSION     AY405612.1  GI:39761586
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
```

```
                     /db_xref="taxon:9606"
     gene            <1. .>933
                     /gene="OR5AR1"
                     /locus_tag="HCM2282"
ORIGIN

Alignment Scores:
Pred. No.:              3.56e-56        Length:         933
Score:                  775.00          Matches:        151
Percent Similarity:     67.64%          Conservative:   58
Best Local Similarity:  48.87%          Mismatches:     100
Query Match:            49.14%          Indels:         0
DB:                     29              Gaps:           0

US-10-023-597-24 (1-311) x AY405612 (1-933)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||      ||||||||||||   ||||||||||||||||      |||:::
Db          1 ATGGATAAAGAAAACAGCTCAATGGTGACTGAGTTTATCTTCATGGGCATCACCCAGGAC 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||   :::::::::      |||   :::||||||      |||   |||   ||||||
Db         61 CCTCAGATGGAGATCATCTTCTTCGTGGTCTTCCTCATAGTTTACCTGGTTAATGTAGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||:::|||:::||||||||||||      :::::::::||||||   ||||||||
Db        121 GGGAATATTGGTATGATTATCCTGATTACAACAGACACTCAGCTTCACACACCCATGTAT 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||      ||||||||   ||||||      :::|||:::   |||   |||:::
Db        181 TTTTTCCTCTGCAACCTCTCCTTTGTTGACCTGGGCTACTCCTCAGCCATTGCCCCCAGG 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||      |||::::::         :::|||||||||:::   |||   :::|||
Db        241 ATGCTGGCTGACTTCCTAACAAATCACAAAGTTATCTCCTTCTCCAGCTGTGCCACCCAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              |||   ||||||      ||||||   :::|||   ::::::|||:::||||||
Db        301 TTTGCTTTTTTTGTAGGTTTTGTGGATGCTGAGTGCTATGTCCTGGCAGCCATGGCCTAT 360

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
                 |||:::|||   ||||||   ||||||   |||:::      ||||||   ||||||
Db        361 GGTCGTTTTGTGGCCATTTGTCGACCCCTCCACTATAGCACCTTCATGTCCAAGCAGGTC 420

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||   |||:::||||||   |||      |||:::      ||||||||||||
Db        421 TGCTTGGCTCTCATGCTGGGCTCTTACCTGGCTGGTCTAGTGAGTTTAGTAGCCCACACT 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
                        |||::::::|||      |||::::::||||||:::   |||:::
Db        481 ACCCTCACCTTCAGCCTGAGTTACTGTGGTTCCAATATCATCAATCATTTCTTCTGCGAA 540

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||   |||||||||   |||||||||:::   :::||||||:::   :::::::::|||
Db        541 ATCCCACCACTCTTGGCCCTCTCTTGCTCAGACACCTACATCAGTGAGATCTTGCTCTTC 600

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
                 :::                        |||:::   :::||||||||||||   |||
Db        601 AGTCTGTGTGGCTTCATTGAATTCAGCACCATCCTCATCATCTTCATCTCCTATACCTTT 660

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||   :::|||:::|||:::   ||||||||||||||   ||||||||||:::
Db        661 ATCCTTGTTGCAATCATCAGAATGCGTTCAGCTGAAGGCCGCCTTAAGGCTTTCTCCACC 720

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||   |||::::::      ::::::||||||:::|||:::      |||   ||||||
Db        721 TGCGGGTCTCACCTTACTGGCATCACCCTCTTCTATGGCACAGTCATGTTTATGTACCTG 780

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              :::|||   |||          |||||||||   |||   :::|||:::|||||||||
Db        781 AGGCCAACATCCAGCTACTCCCTGGACCAAGACAAGTGGGCCTCTGTGTTCTACACGGTT 840

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
                 ::::::||||||   ||||||||||||||||||||||||||||||||||||   |||
Db        841 ATCATCCCCATGTTAAATCCCTTGATCTACAGTTTGCGGAACAAGGATGTGAAAGCTGCT 900

Qy        301 LeuLysArgThrPheSerArgIleSer 309
                 |||:::          :::   |||
Db        901 TTCAAAAAGCTAATTGGAAAAAAATCT 927


Search completed: September 30, 2004, 12:11:38
Job time : 2912 secs
```

This Page Blank (uspto)

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 12:13:16 ; Search time 4423 Seconds
(without alignments)
3047.634 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 311
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934735

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163430_15410/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=olig.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023597_@CGN_1_1_3731_@runat_29092004_163430_15410 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		311	100.0	963	6	AX555964	AX555964 Sequence
2		210	67.5	994	6	AX241500	AX241500 Sequence
3		210	67.5	172991	2	AC083958	AC083958 Homo sapi
4	c	210	67.5	192926	9	AP000916	AP000916 Homo sapi
5		191	61.4	931	6	AX646877	AX646877 Sequence
6		105	33.8	667	9	AY448475	AY448475 Gorilla g
7	c	74	23.8	485	6	AR427386	AR427386 Sequence
8	c	74	23.8	485	6	BD122939	BD122939 EST and e
9		70	22.5	667	9	AY448924	AY448924 Cercopith
10		50	16.1	668	9	AY448357	AY448357 Hylobates
11		41	13.2	667	9	AY448759	AY448759 Cercocebu
12		40	12.9	886	6	AX646875	AX646875 Sequence
13		40	12.9	935	6	AX241493	AX241493 Sequence
14		34	10.9	667	9	AY449293	AY449293 Saimiri s
15		33	10.6	487	6	AX181578	AX181578 Sequence
16		33	10.6	487	9	AF179842	AF179842 Saimiri b
17		28	9.0	668	9	AY448836	AY448836 Callithri
18		23	7.4	667	9	AY448503	AY448503 Alouatta
19		21	6.8	667	9	AY448669	AY448669 Eulemur m
20		21	6.8	667	9	AY449228	AY449228 Ateles fu
21		21	6.8	933	10	AY073346	AY073346 Mus muscu
22		21	6.8	933	10	AY318057	AY318057 Mus muscu
23		21	6.8	235933	2	AC131161	AC131161 Rattus no
24		21	6.8	253120	2	AC097099	AC097099 Rattus no
25	c	21	6.8	258873	2	AC096935	AC096935 Rattus no
26		20	6.4	485	6	AX181444	AX181444 Sequence
27		20	6.4	485	9	AF179764	AF179764 Homo sapi
28		20	6.4	650	9	AF399511	AF399511 Homo sapi
29		20	6.4	930	6	AX242246	AX242246 Sequence
30		20	6.4	930	10	AF282281	AF282281 Mus muscu
31		20	6.4	933	6	AX244609	AX244609 Sequence
32		20	6.4	933	6	AX448443	AX448443 Sequence
33		20	6.4	933	6	BD144473	BD144473 Novel G-p
34		20	6.4	933	10	AY073208	AY073208 Mus muscu
35		20	6.4	933	10	AY318056	AY318056 Mus muscu
36		20	6.4	953	6	AX554459	AX554459 Sequence
37		20	6.4	976	6	AX556206	AX556206 Sequence
38		20	6.4	991	6	AX702842	AX702842 Sequence
39		20	6.4	1333	6	AX646099	AX646099 Sequence
40		20	6.4	1333	9	AB065834	AB065834 Homo sapi
41		20	6.4	1655	6	AX357037	AX357037 Sequence
42	c	20	6.4	150847	9	AP001524	AP001524 Homo sapi
43	c	20	6.4	172991	2	AC083958	AC083958 Homo sapi
44		20	6.4	209420	2	AC109178	AC109178 Mus muscu
45	c	20	6.4	235933	2	AC131161	AC131161 Rattus no

ALIGNMENTS

RESULT 1

```
AX555964
LOCUS       AX555964                963 bp    DNA     linear   PAT 27-NOV-2002
DEFINITION  Sequence 23 from Patent WO0250275.
ACCESSION   AX555964
VERSION     AX555964.1  GI:25899399
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Padigaru,M., Kekuda,R., Li,L., Ballinger,R.A., Casman,S.J.,
            Spytek,K.A., Baumgartner,J.C. and Burgess,C.E.
  TITLE     Novel proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0250275-A 23 27-JUN-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .963
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.37e-316      Length:          963
Score:                    311.00         Matches:         311
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-10-023-597-24 (1-311) x AX555964 (1-963)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 61

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 121

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 181

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 241

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 301

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAG 361

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 421

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 481

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 541

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 601

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 661

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 721

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 781

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 841

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 901

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        902 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 934

RESULT 2
AX241500
LOCUS       AX241500                994 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 248 from Patent WO0127158.
ACCESSION   AX241500
VERSION     AX241500.1  GI:15798375
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 248 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .994
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="(H38g97 nucleotide)"
ORIGIN

Alignment Scores:
Pred. No.:                1.72e-210      Length:          994
Score:                    210.00         Matches:         310
Percent Similarity:       99.36%         Conservative:    0
Best Local Similarity:    99.36%         Mismatches:      1
Query Match:              67.52%         Indels:          2
DB:                       6              Gaps:            0

US-10-023-597-24 (1-311) x AX241500 (1-994)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 240
```

```
Qy          81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 300

Qy         101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTA- 359

Qy         121 -AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVa 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         360 GGACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGT 419

Qy         140 lCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisTh 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         420 GTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATAC 479

Qy         160 rGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAs 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         480 AGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGA 539

Qy         180 pIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePh 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         540 CATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTT 599

Qy         200 eIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPh 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         600 TATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTT 659

Qy         220 eIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSe 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         660 TATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAG 719

Qy         240 rCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLe 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         720 CTGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCT 779

Qy         260 uLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrTh 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         780 CAAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCAC 839

Qy         280 rValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAl 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         840 TGTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGC 899

Qy         300 aLeuLysArgThrPheSerArgIleSerPheSer 311
               |||||||||||||||||||||||||||||||
Db         900 CCTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 933
```

```
RESULT 3
AC083958
LOCUS       AC083958               172991 bp    DNA     linear   HTG 03-MAR-2001
DEFINITION  Homo sapiens clone RP11-18M5, WORKING DRAFT SEQUENCE, 8 unordered
            pieces.
ACCESSION   AC083958
VERSION     AC083958.2  GI:13184080
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 172991)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome, clone RP11-18M5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172991)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
            FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
            Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
            Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G.,
            Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
            O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
            Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 2, 2001 this sequence version replaced gi:10717225.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L11149
                Center clone name: 18_M_5
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 169811 bases at least Q40
                Consensus quality: 170912 bases at least Q30
                Consensus quality: 171602 bases at least Q20
                Insert size: 176000; agarose-fp
                Insert size: 172291; sum-of-contigs
                Quality coverage: 7.3 in Q20 bases; agarose-fp
                Quality coverage: 7.4 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      388: contig of 388 bp in length
            *      389      488: gap of 100 bp
            *      489     3383: contig of 2895 bp in length
            *     3384     3483: gap of 100 bp
            *     3484     6775: contig of 3292 bp in length
            *     6776     6875: gap of 100 bp
            *     6876    69914: contig of 63039 bp in length
            *    69915    70014: gap of 100 bp
            *    70015    87665: contig of 17651 bp in length
            *    87666    87765: gap of 100 bp
            *    87766   112075: contig of 24310 bp in length
            *   112076   112175: gap of 100 bp
            *   112176   166387: contig of 54212 bp in length
            *   166388   166487: gap of 100 bp
            *   166488   172991: contig of 6504 bp in length.
FEATURES             Location/Qualifiers
     source          1. .172991
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-18M5"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .388
                     /note="assembly_fragment
                     clone_end:SP6
```

```
                    vector_side:left"
     misc_feature   489. .3383
                    /note="assembly_fragment"
     misc_feature   3484. .6775
                    /note="assembly_fragment"
     misc_feature   6876. .69914
                    /note="assembly_fragment"
     misc_feature   70015. .87665
                    /note="assembly_fragment"
     misc_feature   87766. .112075
                    /note="assembly_fragment"
     misc_feature   112176. .166387
                    /note="assembly_fragment"
     misc_feature   166488. .172991
                    /note="assembly_fragment
                    clone_end:T7
                    vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:                1.56e-208      Length:          172991
Score:                    210.00         Matches:         310
Percent Similarity:       99.36%         Conservative:    0
Best Local Similarity:    99.36%         Mismatches:      1
Query Match:              67.52%         Indels:          2
DB:                       2              Gaps:            0

US-10-023-597-24 (1-311) x AC083958 (1-172991)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94318 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 94377

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94378 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 94437

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94438 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 94497

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94498 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 94557

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94558 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 94617

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94618 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTA- 94676

Qy        121 -AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVa 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94677 GGACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGT 94736

Qy        140 lCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisTh 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94737 GTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATAC 94796

Qy        160 rGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAs 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94797 AGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGA 94856

Qy        180 pIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePh 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94857 CATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTT 94916

Qy        200 eIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPh 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94917 TATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTT 94976

Qy        220 eIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSe 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94977 TATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAG 95036

Qy        240 rCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLe 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95037 CTGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCT 95096

Qy        260 uLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrTh 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95097 CAAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCAC 95156

Qy        280 rValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAl 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95157 TGTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGC 95216

Qy        300 aLeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db      95217 CCTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 95250

RESULT 4
AP000916/c
LOCUS       AP000916               192926 bp    DNA     linear   PRI 01-FEB-2003
DEFINITION  Homo sapiens genomic DNA, chromosome 11 clone:RP11-728D14, complete
            sequence.
ACCESSION   AP000916
VERSION     AP000916.6  GI:28189528
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA
  JOURNAL   Published Only in Database (1999)
REFERENCE   2  (bases 1 to 192926)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     On Jan 31, 2003 this sequence version replaced gi:15637083.
FEATURES             Location/Qualifiers
     source          1. .192926
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q"
                     /clone="RP11-728D14"
ORIGIN

Alignment Scores:
Pred. No.:                1.72e-208      Length:          192926
Score:                    210.00         Matches:         310
Percent Similarity:       99.36%         Conservative:    0
Best Local Similarity:    99.36%         Mismatches:      1
Query Match:              67.52%         Indels:          2
DB:                       9              Gaps:            0

US-10-023-597-24 (1-311) x AP000916 (1-192926)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102700 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 102641

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     102640 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 102581

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102580 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 102521

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102520 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 102461

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102460 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 102401

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102400 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTA- 102342

Qy        121 -AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVa 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102341 GGACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGT 102282

Qy        140 lCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisTh 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102281 GTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATAC 102222

Qy        160 rGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAs 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102221 AGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGA 102162

Qy        180 pIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePh 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102161 CATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTT 102102

Qy        200 eIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPh 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102101 TATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTT 102042

Qy        220 eIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSe 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102041 TATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAG 101982

Qy        240 rCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLe 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101981 CTGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCT 101922

Qy        260 uLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrTh 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101921 CAAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCAC 101862

Qy        280 rValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAl 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101861 TGTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGC 101802

Qy        300 aLeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db     101801 CCTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 101768
```

RESULT 5
AX646877
LOCUS AX646877 931 bp DNA linear PAT 04-MAR-2003
DEFINITION Sequence 1069 from Patent EP1270724.
ACCESSION AX646877
VERSION AX646877.1 GI:28799404
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.
TITLE Guanosine triphosphate-binding protein coupled receptors
JOURNAL Patent: EP 1270724-A 1069 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology (JP) ; Center for Advanced Science and Technology Incubation, Ltd. (JP)

FEATURES Location/Qualifiers
source 1. .931
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 201. .731
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD69565.1"
/db_xref="GI:28799405"
/db_xref="REMTREMBL:CAD69565"
/translation="MSPQVCLLLLLGVYGMGVFGAVAHTGNIVFLTFCADNLVNHYMC
DILPLLELSCNGSYINVLVIFIVVTVGIGVPIVAVFISYGFILSSILRVSSAEGRSKA
FSSCSSYIIAVSLFFGSGAFTYLKPPSILPLDQGKVSSLFYTTVVPMFNPLIYSLRNK
DVKLALKRTFSRISFS"
ORIGIN

Alignment Scores:

Pred. No.:	1.51e-190	Length:	931
Score:	191.00	Matches:	241
Percent Similarity:	99.18%	Conservative:	0
Best Local Similarity:	99.18%	Mismatches:	1
Query Match:	61.41%	Indels:	2
DB:	6	Gaps:	0

US-10-023-597-24 (1-311) x AX646877 (1-931)

```
Qy         70 AspPheSerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLys 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAG 62

Qy         90 AsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPhe 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCTTT 122

Qy        110 SerGluSerPheIleLeuSerAlaMetValGlu-AspArgTyrValGlyIleCysAsnPr 129
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        123 TCTGAGTCCTTCATCCTGTCGGCGATGGTGTA-GGACCGCTACGTGGGCATCTGTAACCC 181

Qy        129 oLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTy 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTA 241

Qy        149 rGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCy 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTG 301

Qy        169 sAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCy 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTG 361

Qy        189 sAsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyVa 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGGT 421

Qy        209 lProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSe 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GCCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTAG 481

Qy        229 rSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSe 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTTC 541

Qy        249 rLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAs 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGGA 601

Qy        269 pGlnGlyLysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIl 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CCAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAAT 661
```

```
Qy        289 eTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSe 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CTACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAAG 721

Qy        309 rPheSer 311
              |||||||
Db        722 CTTTTCT 728
```

RESULT 6
AY448475
LOCUS AY448475 667 bp DNA linear PRI 07-DEC-2003
DEFINITION Gorilla gorilla clone OLG_68 olfactory receptor-like protein gene, partial sequence.
ACCESSION AY448475
VERSION AY448475.1 GI:38634503
KEYWORDS .
SOURCE Gorilla gorilla (gorilla)
ORGANISM Gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 667)
AUTHORS Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
TITLE Loss of olfactory receptor genes coupled with the acquisition of full trichromatic vision in primates
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 667)
AUTHORS Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-2003) Max Plank Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig 04103, Germany
FEATURES Location/Qualifiers
source 1. .667
/organism="Gorilla gorilla"
/mol_type="genomic DNA"
/db_xref="taxon:9593"
/clone="OLG_68"
/note="sample obtained from Primate Genetics German Primate Center in Goettingen, Germany"
gene <1. .>667
/gene="olfactory receptor-like protein"
ORIGIN

Alignment Scores:

Pred. No.:	2.73e-100	Length:	667
Score:	105.00	Matches:	105
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	33.76%	Indels:	0
DB:	9	Gaps:	0

US-10-023-597-24 (1-311) x AY448475 (1-667)

```
Qy        135 ThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPhe 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 ACCATGTCTCCCCAGGTGTGTTTGCTACTTTTACTGGGTGTCTACGGGATGGGGGTTTTT 274

Qy        155 GlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuVal 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 GGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTC 334

Qy        175 AsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIle 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 AATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATA 394

Qy        195 AsnValLeuValIlePheIleValValThrValGlyIleGlyValProIleValAlaVal 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 AATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTT 454

Qy        215 PheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArg 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 TTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGG 514

Qy        235 SerLysAlaPheSer 239
              |||||||||||||||
Db        515 TCTAAAGCCTTCAGT 529
```

RESULT 7
AR427386/c
LOCUS AR427386 485 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 18883 from patent US 6639063.
ACCESSION AR427386
VERSION AR427386.1 GI:40182496
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 485)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 18883 28-OCT-2003;
FEATURES Location/Qualifiers
source 1. .485
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:

Pred. No.:	7.85e-68	Length:	485
Score:	74.00	Matches:	74
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	23.79%	Indels:	0
DB:	6	Gaps:	0

US-10-023-597-24 (1-311) x AR427386 (1-485)

```
Qy        135 ThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPhe 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 ACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTT 204

Qy        155 GlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuVal 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTC 144

Qy        175 AsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIle 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATA 84

Qy        195 AsnValLeuValIlePheIleValValThrValGlyIleGly 208
              |||||||||||||||||||||||||||||||||||||||
Db         83 AATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGG 42
```

RESULT 8
BD122939/c
LOCUS BD122939 485 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD122939
VERSION BD122939.1 GI:23217884
KEYWORDS JP 2002010789-A/15016.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 485)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 15016 15-JAN-2002;
GENSET CORP
COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/15016
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE PI GIORDANO

```
            PC   C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
             C12N1/21,
            PC   C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
            C12N15/00
            CC   EST and encoded human protein
            FH   Key             Location/Qualifiers
            FT   source          1. .485
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              7.85e-68        Length:         485
Score:                  74.00           Matches:        74
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            23.79%          Indels:         0
DB:                     6               Gaps:           0

US-10-023-597-24 (1-311) x BD122939 (1-485)

Qy        135 ThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPhe 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 ACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTT 204

Qy        155 GlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuVal 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTC 144

Qy        175 AsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIle 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATA 84

Qy        195 AsnValLeuValIlePheIleValValThrValGlyIleGly 208
              |||||||||||||||||||||||||||||||||||||||||
Db         83 AATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGG 42

RESULT 9
AY448924
LOCUS       AY448924                667 bp    DNA     linear   PRI 07-DEC-2003
DEFINITION  Cercopithecus aethiops clone OLG_7 olfactory receptor-like protein
            gene, partial sequence.
ACCESSION   AY448924
VERSION     AY448924.1  GI:38634952
KEYWORDS    .
SOURCE      Cercopithecus aethiops (African green monkey)
  ORGANISM  Cercopithecus aethiops
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Cercopithecus.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
  TITLE     Loss of olfactory receptor genes coupled with the acquisition of
            full trichromatic vision in primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 667)
  AUTHORS   Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2003) Max Plank Institute for Evolutionary
            Anthropology, Deutscher Platz 6, Leipzig 04103, Germany
FEATURES             Location/Qualifiers
     source          1. .667
                     /organism="Cercopithecus aethiops"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9534"
                     /clone="OLG_7"
                     /note="sample obtained from Primate Genetics German
                     Primate Center in Goettingen, Germany"
     gene            <1. .>667
                     /gene="olfactory receptor-like protein"
ORIGIN

Alignment Scores:
Pred. No.:              1.66e-63        Length:         667
Score:                  70.00           Matches:        70
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            22.51%          Indels:         0
DB:                     9               Gaps:           0

US-10-023-597-24 (1-311) x AY448924 (1-667)

Qy        135 ThrMetSerProGlnValCysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPhe 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 ACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTGTTGGGTGTCTATGGGATGGGAGTTTTT 274

Qy        155 GlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuVal 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 GGGGCTGTGGCTCATACAGGAAATATCGTGTTTCTCACCTTTTGTGCAGACAACCTTGTC 334

Qy        175 AsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIle 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 AATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGTAACGGCTCTTACATA 394

Qy        195 AsnValLeuValIlePheIleValValThr 204
              |||||||||||||||||||||||||||||
Db        395 AATGTCCTGGTCATCTTTATTGTTGTGACC 424

RESULT 10
AY448357
LOCUS       AY448357                668 bp    DNA     linear   PRI 07-DEC-2003
DEFINITION  Hylobates syndactylus clone OLG_29 olfactory receptor-like protein
            gene, partial sequence.
ACCESSION   AY448357
VERSION     AY448357.1  GI:38634385
KEYWORDS    .
SOURCE      Hylobates syndactylus (siamang)
  ORGANISM  Hylobates syndactylus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
  TITLE     Loss of olfactory receptor genes coupled with the acquisition of
            full trichromatic vision in primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 668)
  AUTHORS   Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2003) Max Plank Institute for Evolutionary
            Anthropology, Deutscher Platz 6, Leipzig 04103, Germany
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Hylobates syndactylus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9590"
                     /clone="OLG_29"
                     /note="sample obtained from Primate Genetics German
                     Primate Center in Goettingen, Germany"
     gene            <1. .>668
                     /gene="olfactory receptor-like protein"
ORIGIN

Alignment Scores:
Pred. No.:              1.74e-42        Length:         668
Score:                  50.00           Matches:        50
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            16.08%          Indels:         0
DB:                     9               Gaps:           0

US-10-023-597-24 (1-311) x AY448357 (1-668)
```

```
Qy        155 GlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuVal 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 GGGGCTGTGGCTCATACGGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTC 334

Qy        175 AsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIle 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 AATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATA 394

Qy        195 AsnValLeuValIlePheIleValValThr 204
              |||||||||||||||||||||||||||||
Db        395 AATGTCCTGGTCATCTTTATTGTTGTGACC 424

RESULT 11
AY448759
LOCUS       AY448759                667 bp    DNA     linear   PRI 07-DEC-2003
DEFINITION  Cercocebus agilis clone OLG_56 olfactory receptor-like protein
            gene, partial sequence.
ACCESSION   AY448759
VERSION     AY448759.1  GI:38634787
KEYWORDS    .
SOURCE      Cercocebus agilis (agile mangabey)
  ORGANISM  Cercocebus agilis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Cercocebus.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
  TITLE     Loss of olfactory receptor genes coupled with the acquisition of
            full trichromatic vision in primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 667)
  AUTHORS   Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2003) Max Plank Institute for Evolutionary
            Anthropology, Deutscher Platz 6, Leipzig 04103, Germany
FEATURES             Location/Qualifiers
     source          1. .667
                     /organism="Cercocebus agilis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:255237"
                     /clone="OLG_56"
                     /note="sample obtained from Primate Genetics German
                     Primate Center in Goettingen, Germany"
     gene            <1. .>667
                     /gene="olfactory receptor-like protein"
ORIGIN

Alignment Scores:
Pred. No.:              4.98e-33        Length:         667
Score:                  41.00           Matches:        41
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            13.18%          Indels:         0
DB:                     9               Gaps:           0

US-10-023-597-24 (1-311) x AY448759 (1-667)

Qy        164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuPro 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GTGTTTCTTACCTTTTGTGCAGATAACCTTGTCAATCACTACATGTGTGACATCCTTCCC 361

Qy        184 LeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePheIleValVal 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CTTCTTGAGCTCTCCTGTAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTG 421

Qy        204 Thr 204
              |||
Db        422 ACC 424

RESULT 12
AX646875
LOCUS       AX646875                886 bp    DNA     linear   PAT 04-MAR-2003
DEFINITION  Sequence 1067 from Patent EP1270724.
ACCESSION   AX646875
VERSION     AX646875.1  GI:28799398
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.
  TITLE     Guanosine triphosphate-binding protein coupled receptors
  JOURNAL   Patent: EP 1270724-A 1067 02-JAN-2003;
            National Institute of Advanced Industrial Science and Technology
            (JP) ; Center for Advanced Science and Technology Incubation, Ltd.
            (JP)
FEATURES             Location/Qualifiers
     source          1. .886
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             201. .686
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD69564.1"
                     /db_xref="GI:28799399"
                     /db_xref="REMTREMBL:CAD69564"
                     /translation="MPAKNSSSVTAFILSGLTDQPGLQIPAFFLFLGFYAVTVVGNLG
                     LIILIGLNSRLHIPMYFFPFNLSFIDFSYSTTLAPKMLMSFVSENIISYAGCMTQLFF
                     FCFFVFSESYILSAMAYDRYVGICNPLLYTVTMSPQMCLLLLLGVYGMGILGLWLIWE
                     T"
ORIGIN

Alignment Scores:
Pred. No.:              7.18e-32        Length:         886
Score:                  40.00           Matches:        40
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            12.86%          Indels:         0
DB:                     6               Gaps:           0

US-10-023-597-24 (1-311) x AX646875 (1-886)

Qy         28 PhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIle 47
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 TTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAATC 341

Qy         48 LeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSer 67
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTCC 401

RESULT 13
AX241493
LOCUS       AX241493                935 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 241 from Patent WO0127158.
ACCESSION   AX241493
VERSION     AX241493.1  GI:15798368
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE     Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 241 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="(H38g90 nucleotide)"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              7.53e-32        Length:         935
Score:                  40.00           Matches:        40
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            12.86%          Indels:         0
DB:                     6               Gaps:           0

US-10-023-597-24 (1-311) x AX241493 (1-935)

Qy         28 PhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIle 47
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 TTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAATC 141

Qy         48 LeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSer 67
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTCC 201

RESULT 14
AY449293
LOCUS       AY449293                667 bp    DNA     linear   PRI 07-DEC-2003
DEFINITION  Saimiri sciureus clone OLG_66 olfactory receptor-like protein gene,
            partial sequence.
ACCESSION   AY449293
VERSION     AY449293.1  GI:38635321
KEYWORDS    .
SOURCE      Saimiri sciureus (common squirrel monkey)
  ORGANISM  Saimiri sciureus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae;
            Saimiri.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
  TITLE     Loss of olfactory receptor genes coupled with the acquisition of
            full trichromatic vision in primates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 667)
  AUTHORS   Gilad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2003) Max Plank Institute for Evolutionary
            Anthropology, Deutscher Platz 6, Leipzig 04103, Germany
FEATURES             Location/Qualifiers
     source          1. .667
                     /organism="Saimiri sciureus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9521"
                     /clone="OLG_66"
                     /note="sample obtained from Primate Genetics German
                     Primate Center in Goettingen, Germany"
     gene            <1. .>667
                     /gene="olfactory receptor-like protein"
ORIGIN

Alignment Scores:
Pred. No.:              1.13e-25        Length:         667
Score:                  34.00           Matches:        67
Percent Similarity:     97.10%          Conservative:   0
Best Local Similarity:  97.10%          Mismatches:     1
Query Match:            10.93%          Indels:         2
DB:                     9               Gaps:           0

US-10-023-597-24 (1-311) x AY449293 (1-667)

Qy        137 SerProGlnValCysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAla 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 TCTCCTCAGGTGTGCTTGCTCCTTTTGTTGGGTGTCTATGGGATGGGGGTTTTTGGGGCT 280

Qy        157 ValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAsp-AsnLeuValAsnHi 176
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||||
Db        281 GTGGCTCATACAGGAAATATAGTGTTTCTAACCTTTTGTGCAGG-CAACCTTGTCAATCA 339

Qy        176 sTyrMetCysAspIleLeuProLeuGluLeuSerCysAsnGlySerTyrIleAsnVa 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 CTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAATGGCTCTTACATAAATGT 399

Qy        196 lLeuValIlePheIleValValThr 204
              |||||||||||||||||||||||||
Db        400 TCTGGTCATCTTTATTGTTGTGACC 424

RESULT 15
AX181578
LOCUS       AX181578                487 bp    DNA     linear   PAT 07-AUG-2001
DEFINITION  Sequence 366 from Patent WO0146262.
ACCESSION   AX181578
VERSION     AX181578.1  GI:15141666
KEYWORDS    .
SOURCE      Saimiri boliviensis (Bolivian squirrel monkey)
  ORGANISM  Saimiri boliviensis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae;
            Saimiri.
REFERENCE   1
  AUTHORS   Rouquier,S. and Giorgi,D.
  TITLE     Olfactory receptor genes and pseudogenes in primates and mouse
  JOURNAL   Patent: WO 0146262-A 366 28-JUN-2001;
            CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Saimiri boliviensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:27679"
                     /note="Taxon = 27679; gene = SBO222; Accession
                     DDBJ/EMBL/GenBank = AF179"
     CDS             <2. .>487
                     /codon_start=1
                     /product="olfactory receptor"
                     /protein_id="CAC50133.1"
                     /db_xref="GI:15141667"
                     /db_xref="REMTREMBL:CAC50133"
                     /translation="VAICNPLLYMVTMSPQVCLLLLLGVYGMGVLGAVAHTGNIVFLT
                     FCAGNLVNHYMCDILPLLELSCNGSYINVLVIFIVVTIGIGVPIVAIFISYGFILSSN
                     LHISSAEGRSKAFSTCSSHIIAVSLFFGSGAFMYLKPSSVLPLDQGKVSSLFYTIVVP
                     MF"
ORIGIN

Alignment Scores:
Pred. No.:              9.68e-25        Length:         487
Score:                  33.00           Matches:        33
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            10.61%          Indels:         0
DB:                     6               Gaps:           0

US-10-023-597-24 (1-311) x AX181578 (1-487)

Qy        172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGly 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAATGGC 205

Qy        192 SerTyrIleAsnValLeuValIlePheIleValValThr 204
              ||||||||||||||||||||||||||||||||||||||
Db        206 TCTTACATAAATGTTCTGGTCATCTTTATTGTTGTGACC 244


Search completed: September 30, 2004, 15:18:23
Job time : 4546 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 13:24:01 ; Search time 88 Seconds
(without alignments)
1961.247 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 311
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360442

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163431_15438/app_query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=olig.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023597_@CGN_1_1_69_@runat_29092004_163431_15438 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred..No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	74	23.8	485	4	US-09-621-976-18883	Sequence 18883, A
	2	13	4.2	945	4	US-09-016-434-1114	Sequence 1114, Ap
c	3	10	3.2	456	4	US-09-621-976-3066	Sequence 3066, Ap
c	4	10	3.2	561	4	US-09-621-976-3067	Sequence 3067, Ap
	5	10	3.2	966	3	US-08-748-506-8	Sequence 8, Appli
	6	10	3.2	1062	4	US-09-668-680-10	Sequence 10, Appl
	7	10	3.2	1290	2	US-08-827-291A-1	Sequence 1, Appli
	8	10	3.2	1828	3	US-08-988-876-2	Sequence 2, Appli
c	9	9	2.9	457	4	US-09-517-431E-15	Sequence 15, Appl
	10	9	2.9	4989	4	US-09-976-594-114	Sequence 114, App
c	11	8	2.6	360	3	US-08-543-246B-4	Sequence 4, Appli
c	12	8	2.6	405	3	US-08-543-246B-3	Sequence 3, Appli
c	13	8	2.6	405	3	US-08-543-246B-7	Sequence 7, Appli
c	14	8	2.6	419	4	US-09-621-976-8822	Sequence 8822, Ap
	15	8	2.6	472	4	US-09-016-434-1467	Sequence 1467, Ap
c	16	8	2.6	538	4	US-09-621-976-14167	Sequence 14167, A
c	17	8	2.6	645	3	US-08-543-246B-12	Sequence 12, Appl
c	18	8	2.6	693	3	US-08-543-246B-13	Sequence 13, Appl
c	19	8	2.6	699	3	US-08-543-246B-11	Sequence 11, Appl
c	20	8	2.6	1222	3	US-08-543-246B-5	Sequence 5, Appli
c	21	8	2.6	1223	4	US-09-016-434-1347	Sequence 1347, Ap
	22	8	2.6	1297	4	US-09-668-680-11	Sequence 11, Appl
c	23	8	2.6	1333	3	US-08-543-246B-15	Sequence 15, Appl
c	24	8	2.6	1387	3	US-08-543-246B-1	Sequence 1, Appli
	25	8	2.6	1406	4	US-09-149-476-43	Sequence 43, Appl
c	26	8	2.6	1755	3	US-08-543-246B-8	Sequence 8, Appli
	27	8	2.6	1947	4	US-09-668-262A-13	Sequence 13, Appl
c	28	8	2.6	2280	4	US-09-328-352-3524	Sequence 3524, Ap
	29	8	2.6	2688	1	US-08-088-633-3	Sequence 3, Appli
	30	8	2.6	2688	1	US-08-245-756-3	Sequence 3, Appli
	31	8	2.6	2688	1	US-08-441-750-3	Sequence 3, Appli
	32	8	2.6	2688	2	US-08-441-751-3	Sequence 3, Appli
	33	8	2.6	2688	5	PCT-US92-02521-3	Sequence 3, Appli
c	34	8	2.6	2777	3	US-09-282-147-39	Sequence 39, Appl
	35	8	2.6	3033	4	US-09-525-160B-9	Sequence 9, Appli
	36	8	2.6	3213	4	US-09-525-160B-4	Sequence 4, Appli
	37	8	2.6	3459	4	US-09-016-434-1363	Sequence 1363, Ap
c	38	8	2.6	4495	4	US-09-620-312D-352	Sequence 352, App
	39	8	2.6	5690	4	US-09-773-416-13	Sequence 13, Appl
c	40	8	2.6	6008	1	US-07-789-915A-5	Sequence 5, Appli
c	41	8	2.6	6008	1	US-08-005-002C-5	Sequence 5, Appli
c	42	8	2.6	6008	1	US-08-487-203A-5	Sequence 5, Appli
c	43	8	2.6	6755	3	US-08-931-999-4	Sequence 4, Appli
c	44	8	2.6	31063	4	US-09-596-002-20	Sequence 20, Appl
	45	8	2.6	44848	4	US-09-435-739-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-09-621-976-18883/c
; Sequence 18883, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18883
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-18883

Alignment Scores:

Pred. No.:	6.07e-63	Length:	485
Score:	74.00	Matches:	74
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	23.79%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x US-09-621-976-18883 (1-485)

```
Qy        135 ThrMetSerProGlnValCysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPhe 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 ACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTT 204

Qy        155 GlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuVal 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        203 GGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTC 144

Qy        175 AsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIle 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATA 84

Qy        195 AsnValLeuValIlePheIleValValThrValGlyIleGly 208
              |||||||||||||||||||||||||||||||||||||||||
Db         83 AATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGG 42

RESULT 2
US-09-016-434-1114
; Sequence 1114, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1114:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  945 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g1336040
US-09-016-434-1114

Alignment Scores:
Pred. No.:              0.00201         Length:         945
Score:                  13.00           Matches:        13
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.18%           Indels:         0
DB:                     4               Gaps:           0

US-10-023-597-24 (1-311) x US-09-016-434-1114 (1-945)

Qy        286 AsnProLeuIleTyrSerLeuArgAsnLysAspValLys 298
              |||||||||||||||||||||||||||||||||||||||
Db        862 AATCCGTTGATTTATAGTTTGAGAAATAAAGATGTAAAG 900

RESULT 3
US-09-621-976-3066/c
; Sequence 3066, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 3066
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 171..380
US-09-621-976-3066

Alignment Scores:
Pred. No.:              0.812           Length:         456
Score:                  10.00           Matches:        10
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     4               Gaps:           0

US-10-023-597-24 (1-311) x US-09-621-976-3066 (1-456)

Qy        286 AsnProLeuIleTyrSerLeuArgAsnLys 295
              ||||||||||||||||||||||||||||||
Db        424 AATCCTCTAATATATAGTCTCAGGAACAAA 395

RESULT 4
US-09-621-976-3067/c
; Sequence 3067, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 3067
;   LENGTH: 561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 222..392
US-09-621-976-3067

Alignment Scores:
Pred. No.:              0.993           Length:         561
Score:                  10.00           Matches:        10
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     4               Gaps:           0

US-10-023-597-24 (1-311) x US-09-621-976-3067 (1-561)

Qy        286 AsnProLeuIleTyrSerLeuArgAsnLys 295
              ||||||||||||||||||||||||||||||
Db        163 AATCCTCTAATATATAGTCTCAGGAACAAA 134
```

```
RESULT 5
US-08-748-506-8
; Sequence 8, Application US/08748506
; Patent No. 6159707
;  GENERAL INFORMATION:
;    APPLICANT:  Ronnett et al.
;    TITLE OF INVENTION:  NOVEL SPERM RECEPTORS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  US
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,506
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,751
;      FILING DATE:  09-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      REFERENCE/DOCKET NUMBER:  74940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5600
;      TELEFAX:  312-616-5700
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-748-506-8

Alignment Scores:
Pred. No.:              1.69          Length:        966
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.22%         Indels:        0
DB:                     3             Gaps:          0

US-10-023-597-24 (1-311) x US-08-748-506-8 (1-966)

Qy        289 IleTyrSerLeuArgAsnLysAspValLys 298
              |||||||||||||||||||||||||||||||
Db        883 ATCTATAGTTTAAGGAACAAGGATGTCAAG 912

RESULT 6
US-09-668-680-10
; Sequence 10, Application US/09668680
; Patent No. 6436703
; GENERAL INFORMATION:
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION:  No. 6436703el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE:  790CIP2A
;  CURRENT APPLICATION NUMBER: US/09/668,680
;  CURRENT FILING DATE:  2000-09-22
;  PRIOR APPLICATION NUMBER:  09/649,167
;  PRIOR FILING DATE:  2000-08-23
;  PRIOR APPLICATION NUMBER:  09/540,217
;  PRIOR FILING DATE:  2000-03-31
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  pt_FL_genes Version 2.0
; SEQ ID NO 10
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1062)
US-09-668-680-10

Alignment Scores:
Pred. No.:              1.85          Length:        1062
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.22%         Indels:        0
DB:                     4             Gaps:          0

US-10-023-597-24 (1-311) x US-09-668-680-10 (1-1062)

Qy        289 IleTyrSerLeuArgAsnLysAspValLys 298
              |||||||||||||||||||||||||||||||
Db        988 ATCTACAGCCTGAGGAACAAGGATGTGAAG 1017

RESULT 7
US-08-827-291A-1
; Sequence 1, Application US/08827291A
; Patent No. 5874243
;  GENERAL INFORMATION:
;    APPLICANT:  Macina, Roberto
;    APPLICANT:  Sathe, Ganesh
;    TITLE OF INVENTION:  NOVEL OLRCC15 RECEPTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,291A
;      FILING DATE:  28-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GP50001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5015
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1290 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  cDNA
US-08-827-291A-1

Alignment Scores:
Pred. No.:              2.23          Length:        1290
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.22%         Indels:        0
DB:                     2             Gaps:          0

US-10-023-597-24 (1-311) x US-08-827-291A-1 (1-1290)

Qy        286 AsnProLeuIleTyrSerLeuArgAsnLys 295
              |||||||||||||||||||||||||||||
Db       1151 AATCCCCTCATCTACAGCCTCCGCAACAAG 1180

RESULT 8
US-08-988-876-2
; Sequence 2, Application US/08988876
; Patent No. 6063596
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    TITLE OF INVENTION:  G PROTEIN COUPLED RECEPTORS ASSOCIATED
;    TITLE OF INVENTION:  WITH IMMUNE RESPONSE
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/988,876
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0441 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1828 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PROSNOT01
;      CLONE:  364702
US-08-988-876-2

Alignment Scores:
Pred. No.:              3.14          Length:        1828
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.22%         Indels:        0
DB:                     3             Gaps:          0

US-10-023-597-24 (1-311) x US-08-988-876-2 (1-1828)

Qy        221 IleLeuSerSerIleLeuArgValSerSer 230
              |||||||||||||||||||||||||||||
Db       1447 ATTCTTTCCTCCATTCTGAGGGTTTCATCA 1476

RESULT 9
US-09-517-431E-15/c
; Sequence 15, Application US/09517431E
; Patent No. 6593512
; GENERAL INFORMATION:
;  APPLICANT: Vitek, Michael P.
;  APPLICANT:  DAWSON, Hana N.
;  APPLICANT:  LORING, Jeanne F.
;  TITLE OF INVENTION: TRANSGENIC MOUSE EXPRESSING HUMAN TAU GENE
;  FILE REFERENCE: 56816-5002
;  CURRENT APPLICATION NUMBER: US/09/517,431E
;  CURRENT FILING DATE:  2000-03-02
;  PRIOR APPLICATION NUMBER: 60/122,691
;  PRIOR FILING DATE: 1999-03-02
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 457
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-517-431E-15

Alignment Scores:
Pred. No.:              7.61          Length:        457
Score:                  9.00          Matches:       9
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.89%         Indels:        0
DB:                     4             Gaps:          0

US-10-023-597-24 (1-311) x US-09-517-431E-15 (1-457)

Qy         99 SerGlnPhePhePhePheCysPhePhe 107
              |||||||||||||||||||||||||||
Db         43 TCACAGTTCTTTTTCTTTGCTTTTTC 17

RESULT 10
US-09-976-594-114
; Sequence 114, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 114
;   LENGTH: 4989
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 1382924.38
;   NAME/KEY: unsure
;   LOCATION: 99, 544
;   OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-114

Alignment Scores:
Pred. No.:              78.1          Length:        4989
```

```
Score:                  9.00          Matches:       9
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.89%         Indels:        0
DB:                     4             Gaps:          0

US-10-023-597-24 (1-311) x US-09-976-594-114 (1-4989)

Qy        203 ValThrValGlyIleGlyValProIle 211
              |||||||||||||||||||||||||||
Db       1612 GTGACTGTAGGTATCGGTGTTCCTATT 1638

RESULT 11
US-08-543-246B-4/c
; Sequence 4, Application US/08543246B
; Patent No. 6262244
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA and amino acid sequence specific for
;    TITLE OF INVENTION:  natural killer cells
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Michael W. Glynn
;      ADDRESSEE:  No. 6262244artis Corporation
;      STREET:  564 Morris Avenue
;      CITY:  Summit,
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07901-1027
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/543,246B
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/676,663
;      FILING DATE:  28-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/02469
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/122,514
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kassenoff, Melvyn M.
;      REGISTRATION NUMBER:  26,389
;      REFERENCE/DOCKET NUMBER:  118-7704/PCT/CONT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-522-6927
;      TELEFAX:  908-522-6955
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  360 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-543-246B-4

Alignment Scores:
Pred. No.:              56.4          Length:        360
Score:                  8.00          Matches:       8
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.57%         Indels:        0
DB:                     3             Gaps:          0

US-10-023-597-24 (1-311) x US-08-543-246B-4 (1-360)

Qy        124 ValGlyIleCysAsnProLeuLeu 131
              |||||||||||||||||||||||
Db         45 GTTGGAATATGTAATCCACTCCTC 22

RESULT 12
US-08-543-246B-3/c
; Sequence 3, Application US/08543246B
; Patent No. 6262244
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA and amino acid sequence specific for
;    TITLE OF INVENTION:  natural killer cells
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Michael W. Glynn
;      ADDRESSEE:  No. 6262244artis Corporation
;      STREET:  564 Morris Avenue
;      CITY:  Summit,
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07901-1027
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/543,246B
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/676,663
;      FILING DATE:  28-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/02469
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/122,514
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kassenoff, Melvyn M.
;      REGISTRATION NUMBER:  26,389
;      REFERENCE/DOCKET NUMBER:  118-7704/PCT/CONT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-522-6927
;      TELEFAX:  908-522-6955
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  405 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-543-246B-3

Alignment Scores:
Pred. No.:              63.3          Length:        405
Score:                  8.00          Matches:       8
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.57%         Indels:        0
DB:                     3             Gaps:          0

US-10-023-597-24 (1-311) x US-08-543-246B-3 (1-405)

Qy        124 ValGlyIleCysAsnProLeuLeu 131
              |||||||||||||||||||||||
Db         90 GTTGGAATATGTAATCCACTCCTC 67

RESULT 13
US-08-543-246B-7/c
```

```
; Sequence 7, Application US/08543246B
; Patent No. 6262244
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA and amino acid sequence specific for
;    TITLE OF INVENTION:  natural killer cells
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Michael W. Glynn
;      ADDRESSEE:  No. 6262244artis Corporation
;      STREET:  564 Morris Avenue
;      CITY:  Summit,
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07901-1027
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/543,246B
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/676,663
;      FILING DATE:  28-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/02469
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/122,514
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kassenoff, Melvyn M.
;      REGISTRATION NUMBER:  26,389
;      REFERENCE/DOCKET NUMBER:  118-7704/PCT/CONT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-522-6927
;      TELEFAX:  908-522-6955
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  405 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-543-246B-7

Alignment Scores:
Pred. No.:                63.3          Length:          405
Score:                    8.00          Matches:         8
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              2.57%         Indels:          0
DB:                       3             Gaps:            0

US-10-023-597-24 (1-311) x US-08-543-246B-7 (1-405)

Qy        124 ValGlyIleCysAsnProLeuLeu 131
              |||||||||||||||||||||||||
Db         90 GTTGGAATATGTAATCCACTCCTC 67

RESULT 14
US-09-621-976-8822/c
; Sequence 8822, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 8822
;   LENGTH: 419
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-8822

Alignment Scores:
Pred. No.:                65.4          Length:          419
Score:                    8.00          Matches:         8
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              2.57%         Indels:          0
DB:                       4             Gaps:            0

US-10-023-597-24 (1-311) x US-09-621-976-8822 (1-419)

Qy        141 CysLeuLeuLeuLeuLeuGlyVal 148
              ||||||||||||||||||||||||
Db        190 TGCCTCCTGCTGCTGTTGGGAGTG 167

RESULT 15
US-09-016-434-1467
; Sequence 1467, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1467:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  472 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g902329
US-09-016-434-1467
```

Alignment Scores:

Pred. No.:	73.5	Length:	472
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.57%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x US-09-016-434-1467 (1-472)

```
Qy        223 SerSerIleLeuArgValSerSer 230
              ||||||||||||||||||||||||
Db        293 TCCTCCATTCTGAGGGTTTCATCA 316
```

Search completed: September 30, 2004, 16:06:33
Job time : 93 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 13:14:31 ; Search time 2904 Seconds
(without alignments)
3198.050 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 311
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163431_15422/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=olig.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023597_@CGN_1_1_3437_@runat_29092004_163431_15422 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	6.8	335	14	CB173935	CB173935 OR_2o19H0
c 2	21	6.8	475	28	AZ016145	AZ016145 RPCI-23-3
c 3	20	6.4	444	14	CA881328	CA881328 K0989F04-
c 4	20	6.4	472	14	CA880338	CA880338 K0982C03-
5	20	6.4	799	29	AY407110	AY407110 Pan trogl
6	20	6.4	933	29	AY407109	AY407109 Homo sapi
7	20	6.4	933	29	AY407111	AY407111 Mus muscu
8	18	5.8	693	13	BY720370	BY720370 BY720370
9	16	5.1	464	14	CB173917	CB173917 OR_2o19A0
c 10	16	5.1	498	28	AZ721107	AZ721107 RPCI-24-1
11	16	5.1	519	14	CB174259	CB174259 OR_2o39C1
c 12	16	5.1	675	28	AZ430777	AZ430777 1M0215H24
c 13	16	5.1	710	28	BZ874853	BZ874853 CH240_292
c 14	16	5.1	733	28	BZ201201	BZ201201 CH230-303
15	16	5.1	930	29	AY405659	AY405659 Mus muscu
16	15	4.8	527	14	CB173731	CB173731 OR_2o105b
17	15	4.8	578	14	CB174505	CB174505 OR_C10_1_
18	15	4.8	637	14	CB174340	CB174340 OR_2o42A0
c 19	15	4.8	642	29	CE253172	CE253172 tigr-gss-
20	15	4.8	669	28	AZ627033	AZ627033 1M0467B21
21	15	4.8	674	14	CB174506	CB174506 OR_C11_2_
c 22	15	4.8	771	9	AI962273	AI962273 wq46d06.x
23	15	4.8	903	29	AY405609	AY405609 Homo sapi
c 24	13	4.2	321	29	CC568910	CC568910 CH240_443
25	13	4.2	330	28	BZ760015	BZ760015 622_3U2_B
26	13	4.2	346	14	CB173316	CB173316 OR_2o15E0
27	13	4.2	356	28	BH025383	BH025383 RPCI-24-1
28	13	4.2	369	28	BH120437	BH120437 RPCI-24-3
29	13	4.2	386	14	CB172842	CB172842 OR_2o5B11
30	13	4.2	392	14	CB173049	CB173049 OR_2o08B0
31	13	4.2	398	14	CB173176	CB173176 OR_E7_LT0
c 32	13	4.2	424	29	CE056000	CE056000 tigr-gss-
33	13	4.2	432	14	CB174364	CB174364 OR_2o41H0
c 34	13	4.2	454	28	AZ021177	AZ021177 RPCI-23-3
35	13	4.2	458	14	CB174053	CB174053 OR_2o32G0
36	13	4.2	471	29	CE614138	CE614138 tigr-gss-
37	13	4.2	472	14	CB174158	CB174158 OR_2o35C0
38	13	4.2	478	28	AZ734405	AZ734405 RPCI-24-1
c 39	13	4.2	482	29	CC472634	CC472634 CH240_296
40	13	4.2	487	28	AZ632285	AZ632285 1M0486P18
c 41	13	4.2	492	28	BZ224239	BZ224239 CH230-362
42	13	4.2	501	14	CB174092	CB174092 OR_2o33G0
43	13	4.2	503	14	CB173683	CB173683 OR_2o23E0
44	13	4.2	504	14	CB174105	CB174105 OR_2o34C0
45	13	4.2	504	29	CE243689	CE243689 tigr-gss-

ALIGNMENTS

RESULT 1
CB173935
LOCUS CB173935 335 bp mRNA linear EST 09-OCT-2003
DEFINITION OR_2o19H05_E02_010105_13.y3 Adult mouse olfactory epithelium library Mus musculus cDNA clone 2o19H05 5', mRNA sequence.
ACCESSION CB173935
VERSION CB173935.1 GI:37592564
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 335)

```
  AUTHORS   Young,J.M., Shykind,B.M., Lane,R.P., Tonnes-Priddy,L., Ross,J.A.,
            Walker,M., Williams,E.M. and Trask,B.J.
  TITLE     Odorant receptor expressed sequence tags demonstrate olfactory
            expression of over 400 genes, extensive alternate splicing and
            unequal expression levels
  JOURNAL   Genome Biol. 4 (11), R71.1-R71.15 (2003)
COMMENT     Contact: Young JM
            Trask Lab, Division of Human Biology
            Fred Hutchinson Cancer Research Center
            1100 Fairview Avenue N., C3-168, P.O. Box 19024, Seattle, WA
            98109-1024, USA
            Tel: 206 667 1471
            Fax: 206 667 6524
            Email: jayoung@fhcrc.org
            Young gene new name GA_x6K02T2PVTD-31458511-31459443
             Young gene old name GA_x5J8B7W60AJ-1838848-1837916
             Zhang gene name MOR161-4
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .335
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="2o19H05"
                     /tissue_type="Olfactory and respiratory epithelium"
                     /dev_stage="Adult"
                     /clone_lib="Adult mouse olfactory epithelium library"
                     /note="Organ: Olfactory turbinates; Vector:
                     LambdaZAPII-XR; Site_1: EcoRI; Site_2: XhoI; This library
                     was provided by Leslie Vosshall. mRNA was prepared from
                     the olfactory and respiratory epithelium of an adult
                     mouse. Oligo-dT primed cDNA was directionally cloned into
                     Stratagene's lambdaZAPII-XR vector."
ORIGIN

Alignment Scores:
Pred. No.:                  2.71e-07        Length:         335
Score:                      21.00           Matches:        21
Percent Similarity:         100.00%         Conservative:   0
Best Local Similarity:      100.00%         Mismatches:     0
Query Match:                6.75%           Indels:         0
DB:                         14              Gaps:           0

US-10-023-597-24 (1-311) x CB173935 (1-335)

Qy        259 TyrLeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TACCTCAAACCTCCTTCCATTTTGCCCCTTGACCAAGGAAAAGTGTCTTCCTTGTTTTAT 206

Qy        279 Thr 279
              |||
Db        207 ACA 209

RESULT 2
AZ016145/c
LOCUS       AZ016145                475 bp    DNA     linear   GSS 25-FEB-2000
DEFINITION  RPCI-23-379I5.TV RPCI-23 Mus musculus genomic clone RPCI-23-379I5,
            genomic survey sequence.
ACCESSION   AZ016145
VERSION     AZ016145.1  GI:7091529
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 475)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-379I5.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 379  row: I  column: 5
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .475
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-379I5"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

Alignment Scores:
Pred. No.:                  3.4e-07         Length:         475
Score:                      21.00           Matches:        21
Percent Similarity:         100.00%         Conservative:   0
Best Local Similarity:      100.00%         Mismatches:     0
Query Match:                6.75%           Indels:         0
DB:                         28              Gaps:           0

US-10-023-597-24 (1-311) x AZ016145 (1-475)

Qy        259 TyrLeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TACCTCAAACCTCCTTCCATTTTGCCCCTTGACCAAGGAAAAGTGTCTTCCTTGTTTTAT 71

Qy        279 Thr 279
              |||
Db         70 ACA 68

RESULT 3
CA881328/c
LOCUS       CA881328                444 bp    mRNA    linear   EST 20-DEC-2002
DEFINITION  K0989F04-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
            Library (Long) Mus musculus cDNA clone NIA:K0989F04 IMAGE:30092799
            5', mRNA sequence.
ACCESSION   CA881328
VERSION     CA881328.1  GI:27332877
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Piao,Y., Dudekula,D.B., Qian,Y., Martin,P.R., Aiba,K., Vescovi,A.L.
            and Ko,M.S.H.
  TITLE     Systematic Analyses of NIA Mouse Neural Stem Cell
            (Undifferentiated) cDNA Library (Long)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
```

```
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: K0989  row: F  column: 04
            Seq primer: -21M13 Reverse
            High quality sequence stop: 444
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="CD1"
                     /db_xref="niaEST:K0989F04-5N"
                     /db_xref="taxon:10090"
                     /clone="NIA:K0989F04 IMAGE:30092799"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
                     cDNA Library (Long)"
                     /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
                     NotI; Mouse cDNA project by the Laboratory of Genetics,
                     National Institute on Aging (NIA), Intramural Research
                     Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
                     a long-transcript enriched cDNA library (Ref. Genome Res.
                     11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
                     obtained from Dr. Angelo L. Vescovi (Institute for Stem
                     Cell Research, Italy). Double-stranded cDNAs were
                     synthesized with an Oligo(dT) primer [Invitrogen:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from
                     2.0 Microgram of total RNA, treated with T4 DNA
                     polymerase, and purified by ethanol-precipitation. The
                     cDNAs were ligated to Lone-linker LL-Sal4, purified by
                     phenol/chloroform, and separated from free linkers by
                     Centricon 100. Then, the cDNAs were amplified by
                     long-range high fidelity PCR using Ex Taq polymerase
                     (Takara) with a primer Sal4-S. The products were purified
                     by phenol/chloroform and Centricon 100. The cDNAs were
                     digested with SalI and NotI enzymes and cloned into
                     SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
                     coli host was transformed with the ligation mixture by the
                     standard chemical method. The average insert size is about
                     3.8 kb. The library was constructed by Yulan Piao."
ORIGIN

Alignment Scores:
Pred. No.:                2.13e-06        Length:          444
Score:                    20.00           Matches:         20
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              6.43%           Indels:          0
DB:                       14              Gaps:            0

US-10-023-597-24 (1-311) x CA881328 (1-444)

Qy        259 TyrLeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 TACCTCAAACCTCCTTCCATTTTGCCTCTTGACCAAGGAAAAGTGTCTTCCTTGTTCTAT 98

RESULT 4
CA880338/c
LOCUS       CA880338                472 bp    mRNA    linear   EST 20-DEC-2002
DEFINITION  K0982C03-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
            Library (Long) Mus musculus cDNA clone NIA:K0982C03 IMAGE:30092090
            5', mRNA sequence.
ACCESSION   CA880338
VERSION     CA880338.1  GI:27331887
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 472)
  AUTHORS   Piao,Y., Dudekula,D.B., Qian,Y., Martin,P.R., Aiba,K., Vescovi,A.L.
            and Ko,M.S.H.
  TITLE     Systematic Analyses of NIA Mouse Neural Stem Cell
            (Undifferentiated) cDNA Library (Long)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: K0982  row: C  column: 03
            Seq primer: -21M13 Reverse
            High quality sequence stop: 472
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .472
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="CD1"
                     /db_xref="niaEST:K0982C03-5N"
                     /db_xref="taxon:10090"
                     /clone="NIA:K0982C03 IMAGE:30092090"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
                     cDNA Library (Long)"
                     /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
                     NotI; Mouse cDNA project by the Laboratory of Genetics,
                     National Institute on Aging (NIA), Intramural Research
                     Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
                     a long-transcript enriched cDNA library (Ref. Genome Res.
                     11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
                     obtained from Dr. Angelo L. Vescovi (Institute for Stem
                     Cell Research, Italy). Double-stranded cDNAs were
                     synthesized with an Oligo(dT) primer [Invitrogen:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from
                     2.0 Microgram of total RNA, treated with T4 DNA
                     polymerase, and purified by ethanol-precipitation. The
                     cDNAs were ligated to Lone-linker LL-Sal4, purified by
                     phenol/chloroform, and separated from free linkers by
                     Centricon 100. Then, the cDNAs were amplified by
                     long-range high fidelity PCR using Ex Taq polymerase
                     (Takara) with a primer Sal4-S. The products were purified
                     by phenol/chloroform and Centricon 100. The cDNAs were
                     digested with SalI and NotI enzymes and cloned into
                     SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
                     coli host was transformed with the ligation mixture by the
                     standard chemical method. The average insert size is about
                     3.8 kb. The library was constructed by Yulan Piao."
ORIGIN

Alignment Scores:
Pred. No.:                2.22e-06        Length:          472
Score:                    20.00           Matches:         20
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              6.43%           Indels:          0
DB:                       14              Gaps:            0

US-10-023-597-24 (1-311) x CA880338 (1-472)

Qy        259 TyrLeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 TACCTCAAACCTCCTTCCATTTTGCCTCTTGACCAAGGAAAAGTGTCTTCCTTGTTCTAT 98

RESULT 5
AY407110
LOCUS       AY407110                799 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Pan troglodytes OR8B12 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY407110
VERSION     AY407110.1  GI:39763081
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
```

```
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 799)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .799
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>799
                     /gene="OR8B12"
                     /locus_tag="HCM2781"
ORIGIN

Alignment Scores:
Pred. No.:              3.12e-06        Length:         799
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     29              Gaps:           0

US-10-023-597-24 (1-311) x AY407110 (1-799)

Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291

RESULT 6
AY407109
LOCUS       AY407109                933 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Homo sapiens OR8B12 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY407109
VERSION     AY407109.1  GI:39763080
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>933
                     /gene="OR8B12"
                     /locus_tag="HCM2781"
ORIGIN

Alignment Scores:
Pred. No.:              3.45e-06        Length:         933
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     29              Gaps:           0

US-10-023-597-24 (1-311) x AY407109 (1-933)

Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291

RESULT 7
AY407111
LOCUS       AY407111                933 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Mus musculus OR8B12 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY407111
VERSION     AY407111.1  GI:39763082
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 933)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>933
                     /gene="OR8B12"
                     /locus_tag="HCM2781"
ORIGIN

Alignment Scores:
Pred. No.:              3.45e-06        Length:         933
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
```

```
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            6.43%         Indels:        0
DB:                     29            Gaps:          0

US-10-023-597-24 (1-311) x AY407111 (1-933)

Qy        259 TyrLeuLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 TACCTCAAACCTCCTTCCATTTTGCCTCTTGACCAAGGAAAAGTGTCTTCCTTGTTCTAT 831

RESULT 8
BY720370
LOCUS       BY720370                693 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY720370 RIKEN full-length enriched, in vitro fertilized eggs Mus
            musculus cDNA clone 7420497B16 5', mRNA sequence.
ACCESSION   BY720370
VERSION     BY720370.1  GI:27133487
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES              Location/Qualifiers
     source           1. .693
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="7420497B16"
                      /sex="female"
                      /tissue_type="in vitro fertilized eggs"
                      /dev_stage="egg"
                      /lab_host="DH10B"
                      /clone_lib="RIKEN full-length enriched, in vitro
                      fertilized eggs"
                      /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                      prepared and sequenced in Mouse Genome Encyclopedia
                      Project of Genome Exploration Research Group in Riken
                      Genomic Sciences Center and Genome Science Laboratory in
                      RIKEN. Division of Experimental Animal Research in Riken
                      contributed to prepare mouse tissues. 1st strand cDNA was
                      primed with a primer [5'
                      GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3']. cDNA was
                      prepared by using trehalose thermo-activated reverse
                      transcriptase and subsequently enriched for full-length by
                      cap-trapper. Second strand cDNA was prepared with the
                      primer adapter of sequence [5'
                      GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                      was cloned into the XhoI and BamHI sites. Vector: a
                      modified pBluescript KS(+) after bulk excision from Lambda
                      FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
ORIGIN

Alignment Scores:
Pred. No.:              0.000122      Length:        693
Score:                  18.00         Matches:       18
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            5.79%         Indels:        0
DB:                     13            Gaps:          0

US-10-023-597-24 (1-311) x BY720370 (1-693)

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLys 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GTGGTGCCCATGTTCAATCCATTAATCTATAGCCTGAGGAATAAGGATGTCAAA 55

RESULT 9
CB173917
LOCUS       CB173917                464 bp    mRNA    linear   EST 09-OCT-2003
DEFINITION  OR_2o19A03_E01_010105_05.y3 Adult mouse olfactory epithelium
            library Mus musculus cDNA clone 2o19A03 5', mRNA sequence.
ACCESSION   CB173917
VERSION     CB173917.1  GI:37592546
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1  (bases 1 to 464)
  AUTHORS  Young,J.M., Shykind,B.M., Lane,R.P., Tonnes-Priddy,L., Ross,J.A.,
           Walker,M., Williams,E.M. and Trask,B.J.
  TITLE    Odorant receptor expressed sequence tags demonstrate olfactory
           expression of over 400 genes, extensive alternate splicing and
           unequal expression levels
  JOURNAL  Genome Biol. 4 (11), R71.1-R71.15 (2003)
COMMENT    Contact: Young JM
           Trask Lab, Division of Human Biology
           Fred Hutchinson Cancer Research Center
           1100 Fairview Avenue N., C3-168, P.O. Box 19024, Seattle, WA
           98109-1024, USA
           Tel: 206 667 1471
           Fax: 206 667 6524
           Email: jayoung@fhcrc.org
           Young gene new name GA_x6K02T2PVTD-32518237-32519172
            Young gene old name GA_x5J8B7W60AJ-789779-788844
            Zhang gene name MOR161-3
           Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="2o19A03"
                     /tissue_type="Olfactory and respiratory epithelium"
                     /dev_stage="Adult"
                     /clone_lib="Adult mouse olfactory epithelium library"
                     /note="Organ: Olfactory turbinates; Vector:
                     LambdaZAPII-XR; Site_1: EcoRI; Site_2: XhoI; This library
                     was provided by Leslie Vosshall. mRNA was prepared from
                     the olfactory and respiratory epithelium of an adult
                     mouse. Oligo-dT primed cDNA was directionally cloned into
                     Stratagene's lambdaZAPII-XR vector."
ORIGIN

Alignment Scores:
Pred. No.:                0.00402        Length:          464
Score:                    16.00          Matches:         16
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              5.14%          Indels:          0
DB:                       14             Gaps:            0

US-10-023-597-24 (1-311) x CB173917 (1-464)

Qy        174 ValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCys 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GTCAATCATTACATGTGTGACATCCTTCCTCTGCTAGAGCTTTCCTGC 159

RESULT 10
AZ721107/c
LOCUS       AZ721107                498 bp    DNA     linear   GSS 24-JAN-2001
DEFINITION  RPCI-24-156E15.TV RPCI-24 Mus musculus genomic clone
            RPCI-24-156E15, genomic survey sequence.
ACCESSION   AZ721107
VERSION     AZ721107.1  GI:12463469
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-24-156E15.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 156  row: E  column: 15
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .498
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-156E15"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /clone_lib="RPCI-24"
                     /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
                     DNA."
ORIGIN

Alignment Scores:
Pred. No.:                0.00421        Length:          498
Score:                    16.00          Matches:         16
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              5.14%          Indels:          0
DB:                       28             Gaps:            0

US-10-023-597-24 (1-311) x AZ721107 (1-498)

Qy         37 ValThrValValGlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsn 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 GTCACAGTGGTGGGAAACTTGGGCTTGATCATCCTGATTGGCCTCAAT 291

RESULT 11
CB174259
LOCUS       CB174259                519 bp    mRNA    linear   EST 09-OCT-2003
DEFINITION  OR_2o39C11_011004.y1 Adult mouse olfactory epithelium library Mus
            musculus cDNA clone 2o39C11 5', mRNA sequence.
ACCESSION   CB174259
VERSION     CB174259.1  GI:37592888
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Young,J.M., Shykind,B.M., Lane,R.P., Tonnes-Priddy,L., Ross,J.A.,
            Walker,M., Williams,E.M. and Trask,B.J.
  TITLE     Odorant receptor expressed sequence tags demonstrate olfactory
            expression of over 400 genes, extensive alternate splicing and
            unequal expression levels
  JOURNAL   Genome Biol. 4 (11), R71.1-R71.15 (2003)
COMMENT     Contact: Young JM
            Trask Lab, Division of Human Biology
            Fred Hutchinson Cancer Research Center
            1100 Fairview Avenue N., C3-168, P.O. Box 19024, Seattle, WA
            98109-1024, USA
            Tel: 206 667 1471
            Fax: 206 667 6524
            Email: jayoung@fhcrc.org
            Young gene new name GA_x6K02T2Q125-47904535-47903615
             Young gene old name GA_x5J8B7W1Q6G-2523-3443
             Young gene old name GA_x5J8B7W4PJY-571-1491
```

```
            Zhang gene name MOR190-3P
           Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="2o39C11"
                     /tissue_type="Olfactory and respiratory epithelium"
                     /dev_stage="Adult"
                     /clone_lib="Adult mouse olfactory epithelium library"
                     /note="Organ: Olfactory turbinates; Vector:
                     LambdaZAPII-XR; Site_1: EcoRI; Site_2: XhoI; This library
                     was provided by Leslie Vosshall. mRNA was prepared from
                     the olfactory and respiratory epithelium of an adult
                     mouse. Oligo-dT primed cDNA was directionally cloned into
                     Stratagene's lambdaZAPII-XR vector."
ORIGIN

Alignment Scores:
Pred. No.:              0.00432        Length:         519
Score:                  16.00          Matches:        16
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            5.14%          Indels:         0
DB:                     14             Gaps:           0

US-10-023-597-24 (1-311) x CB174259 (1-519)

Qy        283 ProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLys 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 CCCATGTTTAATCCCCTGATCTACAGCTTGAGAAATAAAGATGTAAAA 501

RESULT 12
AZ430777/c
LOCUS       AZ430777                675 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0215H24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0215H24 F, genomic survey sequence.
ACCESSION   AZ430777
VERSION     AZ430777.1  GI:10554790
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0215  row: H  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 675.
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0215H24"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:
Pred. No.:              0.00513        Length:         675
Score:                  16.00          Matches:        16
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            5.14%          Indels:         0
DB:                     28             Gaps:           0

US-10-023-597-24 (1-311) x AZ430777 (1-675)

Qy        286 AsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeu 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AACCCTTTAATCTACAGTTTGAGAAATAAAGATGTTAAACTTGCCCTG 279

RESULT 13
BZ874853/c
LOCUS       BZ874853                710 bp    DNA     linear   GSS 18-MAR-2003
DEFINITION  CH240_292D3.TJ CHORI-240 Bos taurus genomic clone CH240_292D3,
            genomic survey sequence.
ACCESSION   BZ874853
VERSION     BZ874853.1  GI:29102258
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 710)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P.,
            Crawford,A.M. and McEwan,J.C.
  TITLE     Bovine BAC End Sequences from Library CHORI-240
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by AgResearch Ltd., New Zealand and The
            Institute of Genomic Research (TIGR), USA.
            Plate: 292  row: D  column: 3
```

```
           Seq primer: SP6
           Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .710
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_292D3"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

Alignment Scores:
Pred. No.:                0.0053        Length:         710
Score:                    16.00         Matches:        16
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              5.14%         Indels:         0
DB:                       28            Gaps:           0

US-10-023-597-24 (1-311) x BZ874853 (1-710)

Qy        286 AsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeu 301
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        616 AATCCCTTAATTTACAGCTTGAGGAACAAAGATGTTAAGCTTGCTCTG 569

RESULT 14
BZ201201/c
LOCUS       BZ201201                733 bp    DNA     linear   GSS 11-OCT-2002
DEFINITION  CH230-303F3.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
            CH230-303F3, genomic survey sequence.
ACCESSION   BZ201201
VERSION     BZ201201.1  GI:23859253
KEYWORDS    GSS.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 MboI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-303F3.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 303  row: F  column: 3
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .733
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-303F3"
                     /sex="Female"
                     /cell_type="Brain"
                     /clone_lib="CHORI-230 Segment 2"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
ORIGIN

Alignment Scores:
Pred. No.:                0.00541       Length:         733
Score:                    16.00         Matches:        16
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              5.14%         Indels:         0
DB:                       28            Gaps:           0

US-10-023-597-24 (1-311) x BZ201201 (1-733)

Qy        286 AsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeu 301
              ||||||||||||||||||||||||||||||||||||||||||||||
Db         93 AATCCCTTAATTTATAGCTTGAGGAACAAAGATGTTAAACTTGCCTTG 46

RESULT 15
AY405659
LOCUS       AY405659                930 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Mus musculus OR8B4 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY405659
VERSION     AY405659.1  GI:39761633
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>930
                     /gene="OR8B4"
                     /locus_tag="HCM2297"
ORIGIN

Alignment Scores:
Pred. No.:                0.00631       Length:         930
Score:                    16.00         Matches:        16
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              5.14%         Indels:         0
DB:                       29            Gaps:           0

US-10-023-597-24 (1-311) x AY405659 (1-930)
```

```
Qy        286 AsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeu 301
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        853 AACCCACTGATCTATAGTTTGAGGAATAAAGACGTTAAACTTGCCCTA 900
```

Search completed: September 30, 2004, 16:04:59
Job time : 2910 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 30, 2004, 12:11:46 ; Search time 447 Seconds
(without alignments)
2955.681 Million cell updates/sec

Title: US-10-023-597-24
Perfect score: 311
Sequence: 1 MAAENSSSVTEFILAGLIHQ.....LRNKDVKLALKRTFSRISFS 311

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023597/runat_29092004_163430_15400/app_query.fasta_1.455
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=olig.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023597_@CGN_1_1_470_@runat_29092004_163430_15400 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	311	100.0	963	6	AAF88413	Aaf88413 Human GPC
2	210	67.5	994	4	AAH31675	Aah31675 Human olf
3	191	61.4	931	9	ADC86616	Adc86616 Human GPC
4	156	50.2	960	9	ADD12748	Add12748 cDNA enco
5	40	12.9	886	9	ADC86614	Adc86614 Human GPC
6	40	12.9	935	4	AAH31668	Aah31668 Human olf
7	33	10.6	487	4	AAH84099	Aah84099 Mouse olf
8	20	6.4	485	4	AAH84021	Aah84021 Eulemar f
9	20	6.4	930	4	AAH32421	Aah32421 Human olf
10	20	6.4	933	5	AAS15910	Aas15910 DNA encod
11	20	6.4	933	5	AAS42233	Aas42233 Human cDN
12	20	6.4	933	6	ABZ43067	Abz43067 Human GPC
13	20	6.4	933	6	ABK68455	Abk68455 Human DNA
14	20	6.4	933	6	ABK37521	Abk37521 DNA encod
15	20	6.4	953	6	ABT05655	Abt05655 GPCR 4 pr
16	20	6.4	976	6	ABK97207	Abk97207 Human G-p
17	20	6.4	976	9	AAD60328	Aad60328 Human G-p
18	20	6.4	991	6	ABS58782	Abs58782 Human G-p
19	20	6.4	1333	9	ADC85838	Adc85838 Human GPC
20	20	6.4	1655	6	AAD24449	Aad24449 Human G-p
21	19	6.1	486	4	AAH84095	Aah84095 Mouse olf
22	19	6.1	487	4	AAH84089	Aah84089 Saimiri b
23	19	6.1	958	6	AAF88435	Aaf88435 Human GPC
24	19	6.1	958	9	ADD12792	Add12792 cDNA enco
25	19	6.1	993	6	ABQ88089	Abq88089 Human GPC
26	19	6.1	993	9	ADC79407	Adc79407 Human G p
27	18	5.8	487	4	AAH84062	Aah84062 Pongo pyg
28	18	5.8	1015	6	ABQ88091	Abq88091 Human GPC
29	17	5.5	486	4	AAH84022	Aah84022 Eulemar f
30	17	5.5	489	4	AAH84050	Aah84050 Macaca sy
31	17	5.5	871	9	ADC86676	Adc86676 Human GPC
32	17	5.5	937	6	ABS58846	Abs58846 Human G-p
33	17	5.5	937	6	ABQ88360	Abq88360 Human G p
34	17	5.5	973	4	AAH32263	Aah32263 Human olf
35	16	5.1	934	6	AAF88464	Aaf88464 Human GPC
36	16	5.1	934	9	ADD12850	Add12850 cDNA enco
37	16	5.1	953	6	AAF88433	Aaf88433 Human GPC
38	16	5.1	1014	9	ADC79411	Adc79411 Human G p
39	15	4.8	798	4	AAH31604	Aah31604 Human olf
40	15	4.8	801	6	ABZ43085	Abz43085 Human GPC
41	15	4.8	933	4	AAH32246	Aah32246 Human olf
42	15	4.8	933	6	ABK68672	Abk68672 Human DNA
43	15	4.8	936	5	AAS15911	Aas15911 DNA encod
44	15	4.8	936	6	ABK40195	Abk40195 Human G p
45	15	4.8	942	6	ABZ43068	Abz43068 Human GPC

ALIGNMENTS

RESULT 1
AAF88413
ID AAF88413 standard; cDNA; 963 BP.
XX
AC AAF88413;
XX
DT 12-NOV-2002 (first entry)
XX
DE Human GPCRX cDNA SEQ ID 23.
XX
KW Human; anti-HIV; nootropic; antiasthmatic; antiarteriosclerotic; GPCRX;
KW immunosuppressive; immunomodulator; cytostatic; antiinflammatory; AIDS;
KW antidiabetic; neuroprotective; anorectic; haemostatic; antibacterial;
KW fungicide; protozoal; virucide; human G-protein coupled receptor;
KW gene therapy; vaccine; cardiomyopathy; atherosclerosis; diabetes;
KW cell signal processing; cancer; obesity; neurodegenerative disorder;
KW cachexia; anorexia; Alzheimer's disease; Parkinson's disease;
KW immune disorder; graft versus host disease; bronchial asthma;
KW Crohn's disease; multiple sclerosis; haemophilia; infectious disease;
KW idiopathic thrombocytopenic purpura; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200250275-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US048958.
XX
PR 18-DEC-2000; 2000US-0256635P.
PR 21-DEC-2000; 2000US-0257876P.
PR 04-JAN-2001; 2001US-0259743P.

```
PR   10-JAN-2001; 2001US-0260718P.
PR   12-JAN-2001; 2001US-0261498P.
PR   24-JAN-2001; 2001US-0263689P.
PR   08-FEB-2001; 2001US-0267464P.
PR   22-FEB-2001; 2001US-0271021P.
PR   14-MAR-2001; 2001US-0275946P.
PR   23-MAR-2001; 2001US-0278150P.
PR   18-APR-2001; 2001US-0284591P.
PR   23-APR-2001; 2001US-0285718P.
PR   19-JUN-2001; 2001US-0299327P.
PR   16-AUG-2001; 2001US-0312902P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M,  Kekuda R,  Li L,  Ballinger RA,  Casman SJ,  Spytek KA;
PI   Baumgartner JC,  Burgess CE;
XX
DR   WPI; 2002-657419/70.
DR   P-PSDB; AAB71174.
XX
PT   New G-protein coupled receptor polypeptides, useful in gene therapy,
PT   particularly for treating or preventing cardiomyopathy, atherosclerosis,
PT   diabetes, Crohn's disease, hemophilia or cancer in humans.
XX
PS   Claim 3; Page 90; 155pp; English.
XX
CC   This invention describes novel human G-protein coupled receptor (GPCRX)
CC   polypeptides which have anti-HIV, nootropic, antiasthmatic, protozoal;
CC   antiarteriosclerotic; immunosuppressive, immunomodulator, cytostatic,
CC   antiinflammatory, antidiabetic, neuroprotective, anorectic, haemostatic,
CC   antibacterial, fungicide and virucide activity. The products of the
CC   invention can be used in gene therapy or for vaccines. The GPCRX
CC   polypeptide, GPCRX nucleic acid and antibody are useful for treating,
CC   preventing or alleviating a GPCRX-associated disorder or a pathological
CC   state in a subject e.g cardiomyopathy, atherosclerosis, diabetes, or a
CC   disorder related to cell signal processing and metabolic pathway
CC   modulation. The GPCRX polypeptide and nucleic acid are also useful for
CC   diagnosing the presence of or predisposition to a disease associated with
CC   altered levels of GPCRX, particularly cancer. These polypeptides, nucleic
CC   acids and antibodies are also useful for treating or preventing obesity,
CC   neurodegenerative disorders, AIDS, cancer-associated cachexia, anorexia,
CC   Alzheimer's disease, Parkinson's disease, immune disorders, graft versus
CC   host disease, bronchial asthma, Crohn's disease, multiple sclerosis,
CC   haemophilia, idiopathic thrombocytopenic purpura or infectious disease.
CC   They can also be used to screen for potential agonist and antagonist
CC   compounds. The polypeptides are also useful as immunogens to produce
CC   antibodies or as vaccines. Anti-GPCRX antibodies can be used
CC   diagnostically to monitor protein levels in tissue as part of a clinical
CC   testing procedure such as in determining the efficacy of a given
CC   treatment regimen. The host cells are useful in producing non-human
CC   transgenic animals which are useful for studying the function and/or
CC   activity of GPCRX protein and for identifying and/or evaluating
CC   modulators of GPCRX protein activity. AAF88402-AAF88465 encode the human
CC   GPCRX proteins represented in AAB71163-AAB71226
XX
SQ   Sequence 963 BP; 193 A; 247 C; 204 G; 319 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.41e-288      Length:          963
Score:                    311.00         Matches:         311
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-10-023-597-24 (1-311) x AAF88413 (1-963)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 61

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         62 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 121

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        122 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 181

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        182 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 241

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        242 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 301

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        302 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAG 361

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        362 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 421

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        422 TGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 481

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        482 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGAC 541

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        542 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 601

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        602 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 661

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        662 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 721

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        722 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 781

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        782 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 841

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        842 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 901

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        902 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 934

RESULT 2
AAH31675
ID   AAH31675 standard; DNA; 994 BP.
XX
AC   AAH31675;
XX
DT   30-JUL-2001  (first entry)
XX
DE   Human olfactory receptor polynucleotide, SEQ ID NO: 248.
XX
KW   Human; olfactory receptor; OR; primary scent determination;
KW   secondary scent determination; polypeptide library; odour receptor;
KW   scent profile; scent fingerprint; scent representation; ds.
XX
OS   Homo sapiens.
```

XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027582.
XX
PR 08-OCT-1999; 99US-0158615P.
PR 24-FEB-2000; 2000US-0184809P.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists.
XX
PS Claim 8; Page 276; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides which
CC encode polypeptides involved in olfactory sensation. The polynucleotides
CC can be used in screening for olfactory agonists and antagonists. The
CC methods allow for the determination of primary scents and the
CC identification of the odour receptors used to detect these primary
CC scents. The methods also enable determination of secondary scents and the
CC identification of combinations of odour receptors that are involved in
CC detecting such secondary scents. This enables the construction of a scent
CC representation (also called a scent fingerprint or scent profile), which
CC may be used to re-create and edit scents. Libraries of olfactory
CC receptors are useful for determining the interaction pattern of a
CC composition with the receptors, and can be used for determining
CC differences in the olfactory faculties of different individuals
XX
SQ Sequence 994 BP; 201 A; 250 C; 207 G; 336 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.21e-191	Length:	994
Score:	210.00	Matches:	310
Percent Similarity:	99.36%	Conservative:	0
Best Local Similarity:	99.36%	Mismatches:	1
Query Match:	67.52%	Indels:	2
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x AAH31675 (1-994)

```
Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 60

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 120

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 180

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 240

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 300

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGTA- 359

Qy        121 -AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVa 140
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GGACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGT 419

Qy        140 lCysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisTh 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GTGTTTGCTCCTTTTACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGCTCATAC 479

Qy        160 rGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAs 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AGGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACATGTGTGA 539

Qy        180 pIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePh 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTT 599

Qy        200 eIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPh 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTT 659

Qy        220 eIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSe 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAG 719

Qy        240 rCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLe 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CTGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCT 779

Qy        260 uLysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrTh 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 CAAACCCCCTTCCATTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCAC 839

Qy        280 rValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAl 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TGTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGC 899

Qy        300 aLeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        900 CCTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 933
```

RESULT 3
ADC86616
ID ADC86616 standard; DNA; 931 BP.
XX
AC ADC86616;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human GPCR gene SEQ ID NO:1069.
XX
KW ds; gene; human; GPCR;
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1270724-A2.
XX
PD 02-JAN-2003.
XX
PF 18-JUN-2002; 2002EP-00013517.
XX
PR 18-JUN-2001; 2001JP-00246789.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX
DR WPI; 2003-315783/31.
DR P-PSDB; ADC86617.
XX
PT New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.

```
XX
PS   Claim 1; SEQ ID NO 1069; 28pp; English.
XX
CC   The invention relates to a novel polynucleotide encoding a guanosine
CC   triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC   the invention may have a use in gene therapy. The polynucleotide and
CC   polypeptide are useful for preparing a composition for treating a patient
CC   in need of increased or suppressed activity or expression of the
CC   guanosine triphosphate-binding protein coupled receptor. The
CC   polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC   invention.
XX
SQ   Sequence 931 BP; 202 A; 215 C; 180 G; 334 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.36e-173       Length:         931
Score:                  191.00          Matches:        241
Percent Similarity:     99.18%          Conservative:   0
Best Local Similarity:  99.18%          Mismatches:     1
Query Match:            61.41%          Indels:         2
DB:                     9               Gaps:           0

US-10-023-597-24 (1-311) x ADC86616 (1-931)

Qy         70 AspPheSerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLys 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GATTTTAGTTTCTCTACGACCATCATTCCCAAAATGCTGATGAGTTTTGTCTCAAGGAAG 62

Qy         90 AsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPhe 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AACATTATTTCCTTCACAGGGTGTATGAGTCAGTTCTTCTTCTTCTGTTTCTTTGTCTTT 122

Qy        110 SerGluSerPheIleLeuSerAlaMetValGlu-AspArgTyrValGlyIleCysAsnPr 129
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        123 TCTGAGTCCTTCATCCTGTCGGCGATGGTGTA-GGACCGCTACGTGGGCATCTGTAACCC 181

Qy        129 oLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuGlyValTy 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACTGTTGTACACGATCACCATGTCTCCCCAGGTGTGTTTGCTCCTTTTACTGGGTGTCTA 241

Qy        149 rGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCy 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGGGATGGGGGTTTTTGGGGCTGTGGCTCATACAGGAAATATAGTGTTTCTCACCTTTTG 301

Qy        169 sAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCy 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGCAGACAACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTG 361

Qy        189 sAsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyVa 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAACGGCTCTTACATAAATGTCCTGGTCATCTTTATTGTTGTGACCGTTGGCATTGGGGT 421

Qy        209 lProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSe 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GCCCATTGTTGCCGTTTTTATCTCTTATGGTTTTATTCTTTCCAGCATTCTCCGCGTTAG 481

Qy        229 rSerAlaGluGlyArgSerLysAlaPheSerSerCysSerSerTyrIleIleAlaValSe 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGCTGCAGCTCCTACATAATTGCAGTTTC 541

Qy        249 rLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAs 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTCAAACCCCCTTCCATTTTACCCCTGGA 601

Qy        269 pGlnGlyLysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIl 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CCAGGGGAAAGTGTCCTCCCTGTTCTATACCACTGTGGTGCCCATGTTTAACCCATTAAT 661

Qy        289 eTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSe 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CTACAGCCTGAGGAATAAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGAATAAG 721

Qy        309 rPheSer 311
```

```
              |||||||
Db        722 CTTTTCT 728

RESULT 4
ADD12748
ID   ADD12748 standard; cDNA; 960 BP.
XX
AC   ADD12748;
XX
DT   01-JAN-2004  (first entry)
XX
DE   cDNA encoding novel human olfactory receptor (OR) seq id 12.
XX
KW   cardiant; antiarteriosclerotic; antidiabetic;
KW   G-Protein Coupled Receptor modulator; gene therapy; olfactory receptor;
KW   G-Protein Coupled Receptor X-associated disorder;
KW   GPCRX associated disorder; cardiomyopathy; atherosclerosis;
KW   cell signal processing; metabolic pathway modulation; cancer; diabetes;
KW   human; G-protein coupled receptor; GPCR; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003109692-A1.
XX
PD   12-JUN-2003.
XX
PF   18-DEC-2001; 2001US-00023597.
XX
PR   18-DEC-2000; 2000US-0256635P.
PR   21-DEC-2000; 2000US-0257876P.
PR   04-JAN-2001; 2001US-0259743P.
PR   10-JAN-2001; 2001US-0260718P.
PR   12-JAN-2001; 2001US-0261498P.
PR   24-JAN-2001; 2001US-0263689P.
PR   08-FEB-2001; 2001US-0267464P.
PR   22-FEB-2001; 2001US-0271021P.
PR   14-MAR-2001; 2001US-0275946P.
PR   23-MAR-2001; 2001US-0278150P.
PR   18-APR-2001; 2001US-0284591P.
PR   23-APR-2001; 2001US-0285718P.
PR   19-JUN-2001; 2001US-0299327P.
PR   16-AUG-2001; 2001US-0312902P.
XX
PA   (PADI/) PADIGARU M.
PA   (KEKU/) KEKUDA R.
PA   (LILL/) LI L.
PA   (BALL/) BALLINGER R A.
PA   (CASM/) CASMAN S J.
PA   (SPYT/) SPYTEK K A.
PA   (BAUM/) BAUMGARTNER J C.
PA   (BURG/) BURGESS C E.
XX
PI   Padigaru M,  Kekuda R,  Li L,  Ballinger RA,  Casman SJ,  Spytek KA;
PI   Baumgartner JC,  Burgess CE;
XX
DR   WPI; 2003-801280/75.
DR   P-PSDB; ADD12749.
XX
PT   New isolated olfactory receptor polypeptide for treating or preventing a
PT   G-Protein Coupled Receptor associated disorder e.g. cardiomyopathy or
PT   atherosclerosis.
XX
PS   Claim 8; SEQ ID NO 23; 70pp; English.
XX
CC   The invention describes an isolated olfactory receptor polypeptide (I).
CC   (I) is used to identify an agent that binds to it by contacting (I) with
CC   the agent and determining binding. The agent is a cellular receptor or a
CC   downstream effector. (I) Is also used to identify an agent that modulates
CC   the expression or activity of (I) by contacting a cell expressing (I)
CC   with the agent and determining whether the agent modulates expression or
CC   activity of (I). (I), or nucleic acid encoding (I), is used to treat or
CC   prevent a G-Protein Coupled Receptor (GPCR)X-associated disorder in a
CC   subject, preferably a human, such as cardiomyopathy or atherosclerosis,
```

```
CC   or a disorder related to cell signal processing or metabolic pathway
CC   modulation. (I) Is used to determine the presence or predisposition to a
CC   disease associated with altered levels of (I), preferably cancer, in a
CC   mammal. A nucleic acid encoding (I) is used to determine the presence or
CC   predisposition to a disease associated with altered levels of the nucleic
CC   acid, preferably cancer, in a mammalian subject. An antibody that binds
CC   to (I) is used to treat or prevent a GPCRX-associated disorder, such as
CC   diabetes or a disorder related to cell signal processing and metabolic
CC   pathway modulation. An antibody that specifically binds (I) is used to
CC   determine the presence of amount of (I) in a sample by contacting the
CC   sample with the antibody and determining the presence or amount of bound
CC   antibody. This sequence encodes a novel human olfactory receptor (OR), a
CC   G protein coupled receptor (GPCR).
XX
SQ   Sequence 960 BP; 193 A; 247 C; 202 G; 318 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.21e-139      Length:         960
Score:                  156.00         Matches:        309
Percent Similarity:     99.36%         Conservative:   0
Best Local Similarity:  99.36%         Mismatches:     2
Query Match:            50.16%         Indels:         2
DB:                     9              Gaps:           0

US-10-023-597-24 (1-311) x ADD12748 (1-960)

Qy          1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGCTGCCGAGAACTCCTCCTCCGTGACAGAGTTTATCCTCGCAGGCTTAATCCACCAG 61

Qy         21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCGGGACTCCAGGTCCCCGTCTTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTG 121

Qy         41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGAACCTGGGCTTGATAATCCTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTAC 181

Qy         61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TTTTTCCCCTTCAACTTGTCCCTCGTAGATTTTAGTTTCTCTACGACCATCATTCCCAAA 241

Qy         81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGCTGATGAGTTTTGTCTCAAGGAAGAACATTATTTCCTTCACAGGGTGTATGAGTCAG 301

Qy        101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTCTTCTTCTTCTGTTTCTTTGTCTTTTCTGAGTCCTTCATCCTGTCGGCGATGGTGGAG 361

Qy        121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GACCGCTACGTGGGCATCTGTAACCCACTGTTGTACACGATCACCATGTCTCCCCAGGTG 421

Qy        141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
              |||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db        422 TGTTTGCTCCTTTTACTGGGTGTCTACGGGAT-GGGGTTTTTGGGGCTGTGGCTCATACA 480

Qy        161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGAAATATAGTGTTTCTCACCTTTTGTGCAGACAACCTTGTCAATCACTACA--TGTGAC 538

Qy        181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 ATCCTTCCCCTTCTTGAGCTCTCCTGCAACGGCTCTTACATAAATGTCCTGGTCATCTTT 598

Qy        201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 ATTGTTGTGACCGTTGGCATTGGGGTGCCCATTGTTGCCGTTTTTATCTCTTATGGTTTT 658

Qy        221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 ATTCTTTCCAGCATTCTCCGCGTTAGTTCTGCTGAGGGCAGGTCTAAAGCCTTCAGTAGC 718

Qy        241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TGCAGCTCCTACATAATTGCAGTTTCTCTTTTCTTTGGGTCAGGAGCTTTTACGTACCTC 778

Qy        261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 AAACCCCCTTCCATTTTACCCCTGGACCAGGGGAAAGTGTCCTCCCTGTTCTATACCACT 838

Qy        281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 GTGGTGCCCATGTTTAACCCATTAATCTACAGCCTGAGGAATAAGGATGTCAAACTTGCC 898

Qy        301 LeuLysArgThrPheSerArgIleSerPheSer 311
              |||||||||||||||||||||||||||||||
Db        899 CTGAAGAGAACCTTTTCCAGAATAAGCTTTTCT 931

RESULT 5
ADC86614
ID   ADC86614 standard; DNA; 886 BP.
XX
AC   ADC86614;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human GPCR gene SEQ ID NO:1067.
XX
KW   ds; gene; human; GPCR;
KW   guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1270724-A2.
XX
PD   02-JAN-2003.
XX
PF   18-JUN-2002; 2002EP-00013517.
XX
PR   18-JUN-2001; 2001JP-00246789.
XX
PA   (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Suwa M,  Asai K,  Akiyama Y,  Aburatani H;
XX
DR   WPI; 2003-315783/31.
DR   P-PSDB; ADC86615.
XX
PT   New polynucleotide, useful for preparing a composition for treating a
PT   patient in need of increased or suppressed activity or expression of the
PT   guanosine triphosphate-binding protein coupled receptor.
XX
PS   Claim 1; SEQ ID NO 1067; 28pp; English.
XX
CC   The invention relates to a novel polynucleotide encoding a guanosine
CC   triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC   the invention may have a use in gene therapy. The polynucleotide and
CC   polypeptide are useful for preparing a composition for treating a patient
CC   in need of increased or suppressed activity or expression of the
CC   guanosine triphosphate-binding protein coupled receptor. The
CC   polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC   invention.
XX
SQ   Sequence 886 BP; 168 A; 233 C; 176 G; 309 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.67e-28       Length:         886
Score:                  40.00          Matches:        40
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            12.86%         Indels:         0
DB:                     9              Gaps:           0
```

US-10-023-597-24 (1-311) x ADC86614 (1-886)

```
Qy         28 PhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIle 47
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 TTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAATC 341

Qy         48 LeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSer 67
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTCC 401
```

RESULT 6
AAH31668
ID AAH31668 standard; DNA; 935 BP.
XX
AC AAH31668;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 241.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027582.
XX
PR 08-OCT-1999; 99US-0158615P.
PR 24-FEB-2000; 2000US-0184809P.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists.
XX
PS Claim 8; Page 273; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides which
CC encode polypeptides involved in olfactory sensation. The polynucleotides
CC can be used in screening for olfactory agonists and antagonists. The
CC methods allow for the determination of primary scents and the
CC identification of the odour receptors used to detect these primary
CC scents. The methods also enable determination of secondary scents and the
CC identification of combinations of odour receptors that are involved in
CC detecting such secondary scents. This enables the construction of a scent
CC representation (also called a scent fingerprint or scent profile), which
CC may be used to re-create and edit scents. Libraries of olfactory
CC receptors are useful for determining the interaction pattern of a
CC composition with the receptors, and can be used for determining
CC differences in the olfactory faculties of different individuals
XX
SQ Sequence 935 BP; 189 A; 238 C; 190 G; 318 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.76e-28	Length:	935
Score:	40.00	Matches:	40
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.86%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x AAH31668 (1-935)

```
Qy         28 PhePheLeuPheLeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIle 47
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 TTCTTCCTGTTTCTAGGTTTCTACGCGGTCACGGTGGTGGGGAACCTGGGCTTGATAATC 141

Qy         48 LeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSer 67
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CTGATAGGGCTCAACTCTCGCCTGCATATCCCCATGTACTTTTTCCCCTTCAACTTGTCC 201
```

RESULT 7
AAH84099
ID AAH84099 standard; DNA; 487 BP.
XX
AC AAH84099;
XX
DT 25-SEP-2001 (first entry)
XX
DE Mouse olfactory receptor encoding gene 6.
XX
KW Olfactory receptor; primate; mouse; human; food processing industry;
KW aromas; perfumery; toxic substance; ds.
XX
OS Mus musculus domesticus.
XX
PN WO200146262-A2.
XX
PD 28-JUN-2001.
XX
PF 22-DEC-2000; 2000WO-IB002017.
XX
PR 22-DEC-1999; 99US-0171746P.
PR 21-DEC-2000; 2000US-00747155.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Rouquier S, Giorgi D;
XX
DR WPI; 2001-381911/40.
DR P-PSDB; AAG98584.
XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful for
PT analysis odors e.g. in food processing and perfumery.
XX
PS Claim 1; Page 415-416; 482pp; English.
XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours)
XX
SQ Sequence 487 BP; 86 A; 120 C; 105 G; 176 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.78e-22	Length:	487
Score:	33.00	Matches:	33
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.61%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x AAH84099 (1-487)

```
Qy        172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsnGly 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AACCTTGTCAATCACTACATGTGTGACATCCTTCCCCTTCTTGAGCTCTCCTGCAATGGC 205

Qy        192 SerTyrIleAsnValLeuValIlePheIleValValThr 204
              |||||||||||||||||||||||||||||||||||||||
Db        206 TCTTACATAAATGTTCTGGTCATCTTTATTGTTGTGACC 244
```

RESULT 8

AAH84021
ID AAH84021 standard; DNA; 485 BP.
XX
AC AAH84021;
XX
DT 06-AUG-2003 (revised)
DT 25-SEP-2001 (first entry)
XX
DE Eulemar fulvus olfactory receptor encoding gene 5.
XX
KW Olfactory receptor; primate; mouse; human; food processing industry;
KW aromas; perfumery; toxic substance; ds.
XX
OS Eulemur fulvus.
XX
PN WO200146262-A2.
XX
PD 28-JUN-2001.
XX
PF 22-DEC-2000; 2000WO-IB002017.
XX
PR 22-DEC-1999; 99US-0171746P.
PR 21-DEC-2000; 2000US-00747155.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Rouquier S, Giorgi D;
XX
DR WPI; 2001-381911/40.
DR P-PSDB; AAG98524.
XX
PT Nucleic acids encoding primate and murine olfactory receptors, useful for
PT analysis odors e.g. in food processing and perfumery.
XX
PS Claim 1; Page 280-281; 482pp; English.
XX
CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10
CC primate species, mouse and human. The nucleic acids and receptors may be
CC used in the food processing industry (e.g. for the detection of aromas,
CC quality control and sample analysis), in perfumery (e.g. for the analysis
CC or comparison of perfumes) and in the environment (e.g. for the detection
CC of toxic substances and/or trapping of odours). (Updated on 06-AUG-2003
CC to correct OS field.)
XX
SQ Sequence 485 BP; 90 A; 116 C; 106 G; 173 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.37e-09	Length:	485
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x AAH84021 (1-485)

```
Qy        171 AspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsn 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GACAACCTTGTCAATCACTATATGTGTGACATCCTTCCTCTCCTTGAGCTCTCCTGCAAC 201
```

RESULT 9
AAH32421
ID AAH32421 standard; DNA; 930 BP.
XX
AC AAH32421;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 994.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027582.
XX
PR 08-OCT-1999; 99US-0158615P.
PR 24-FEB-2000; 2000US-0184809P.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists.
XX
PS Claim 8; Page 577; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides which
CC encode polypeptides involved in olfactory sensation. The polynucleotides
CC can be used in screening for olfactory agonists and antagonists. The
CC methods allow for the determination of primary scents and the
CC identification of the odour receptors used to detect these primary
CC scents. The methods also enable determination of secondary scents and the
CC identification of combinations of odour receptors that are involved in
CC detecting such secondary scents. This enables the construction of a scent
CC representation (also called a scent fingerprint or scent profile), which
CC may be used to re-create and edit scents. Libraries of olfactory
CC receptors are useful for determining the interaction pattern of a
CC composition with the receptors, and can be used for determining
CC differences in the olfactory faculties of different individuals
XX
SQ Sequence 930 BP; 190 A; 251 C; 188 G; 301 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.58e-09	Length:	930
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x AAH32421 (1-930)

```
Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291
```

RESULT 10
AAS15910
ID AAS15910 standard; cDNA; 933 BP.
XX
AC AAS15910;
XX
DT 25-JAN-2002 (first entry)
XX
DE DNA encoding G-protein coupled receptor (GCREC) #14.
XX
KW G-protein coupled receptor; GCREC; vaccine; gene therapy;
KW cell proliferation disorder; cancer; arteriosclerosis;
KW neurological disorder; epilepsy; stroke; cardiovascular disorder;
KW hypertension; ischaemic heart disease; gastrointestinal disorder;
KW anorexia; peptic ulcer; autoimmune disorder; inflammatory disorder;
KW diabetes; osteoporosis; psoriasis; metabolic disorder; obesity;
KW schizophrenic disorder; neuroskeletal disorder; ss.
XX

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .933
FT                   /*tag= a
FT                   /product= "GCREC 14"
FT                   /note= "G-protein coupled receptor 14"
XX
PN   WO200166742-A2.
XX
PD   13-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US006814.
XX
PR   03-MAR-2000; 2000US-0186854P.
PR   10-MAR-2000; 2000US-0188384P.
PR   17-MAR-2000; 2000US-0190453P.
PR   20-MAR-2000; 2000US-0190730P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lal P, Tang YT, Patterson C, Yao MG, Shih LL, Tribouley CM;
PI   Lu DAM, Yue H, Khan FA, Policky JL, Au-Young J, Yang J, Harland L;
PI   Walsh RT, Lo TP, Borowsky ML;
XX
DR   WPI; 2001-656776/75.
DR   P-PSDB; AAU10313.
XX
PT   Novel G-protein coupled receptor polypeptides, for treating and
PT   preventing autoimmune/inflammatory disorders, neurological disorders,
PT   cell proliferative disorders, cardiovascular disorders and viral
PT   infections.
XX
PS   Claim 5; Page 138; 141pp; English.
XX
CC   The invention describes a novel isolated polypeptide, selected from a
CC   group of 21 G-protein coupled receptor polypeptides (GCREC) and useful in
CC   vaccines and gene therapy. The polypeptide (I) is useful for screening
CC   for agonist or antagonist of (I), compounds specifically binding to (I),
CC   or compounds that modulate the activity of (I). The polynucleotide
CC   encoding (1) is useful for screening a compound for effectiveness in
CC   altering expression of a target polynucleotide comprising (II), by
CC   exposing a sample comprising the target polynucleotide to a compound,
CC   detecting altered expression of the target polynucleotide, and comparing
CC   the expression of the target polynucleotide in the presence of varying
CC   amounts of compound and in the absence of the compound. (I) and (II) are
CC   useful for diagnosis, treatment and prevention of cell proliferative
CC   disorders (e.g. cancers, arteriosclerosis, atherosclerosis), neurological
CC   disorders (e.g. epilepsy, stroke, schizophrenic disorders and
CC   neuroskeletal disorders), cardiovascular disorders (e.g. hypertension,
CC   ischaemic heart disease), gastrointestinal disorders (e.g. anorexia,
CC   peptic ulcer), autoimmune/inflammatory disorders (e.g. diabetes mellitus,
CC   osteoporosis, psoriasis), and metabolic disorders such as obesity.
CC   Furthermore, the polynucleotide is useful; as primers for detecting
CC   single nucleotide polymorphisms; as elements in microarray, to monitor or
CC   measure protein-protein interactions, drug-target interactions, and gene
CC   expression profiles; to generate a transcript image of a tissue or cell
CC   type, and to generate hybridisation probes useful in mapping the
CC   naturally occurring genomic sequence. This sequence encodes G-protein
CC   coupled receptor 14, one of 21 GCREC proteins described in the method of
CC   the invention
XX
SQ   Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.58e-09      Length:          933
Score:                  20.00         Matches:         20
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            6.43%         Indels:          0
DB:                     5             Gaps:            0

US-10-023-597-24 (1-311) x AAS15910 (1-933)
```

```
Qy          79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291

RESULT 11
AAS42233
ID   AAS42233 standard; cDNA; 933 BP.
XX
AC   AAS42233;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human cDNA encoding olfactory receptor AOLFR26.
XX
KW   Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW   ss; food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS   Homo sapiens.
XX
PN   WO200168805-A2.
XX
PD   20-SEP-2001.
XX
PF   13-MAR-2001; 2001WO-US007771.
XX
PR   13-MAR-2000; 2000US-0188914P.
PR   24-MAR-2000; 2000US-0192033P.
PR   12-APR-2000; 2000US-0198474P.
PR   24-APR-2000; 2000US-0199335P.
PR   26-MAY-2000; 2000US-0207702P.
PR   23-JUN-2000; 2000US-0213849P.
PR   16-AUG-2000; 2000US-0226534P.
PR   07-SEP-2000; 2000US-0230732P.
PR   07-FEB-2001; 2001US-0266862P.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Zozulya S;
XX
DR   WPI; 2001-570867/64.
DR   P-PSDB; AAU24540.
XX
PT   Nucleic acids encoding human olfactory G protein-coupled receptors,
PT   useful for screening for compounds involved in olfactory sensation, where
PT   the compounds can be used in the food, pharmaceutical and cosmetic
PT   industries to customize odors.
XX
PS   Claim 1; Page 94; 319pp; English.
XX
CC   The invention relates to nucleic acids encoding human olfactory
CC   receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC   specifically recognise molecules, odourants, that elicit specific
CC   olfactory sensation. The human olfactory receptors and polynucleotides
CC   encoding them are useful for screening a library of chemical compounds
CC   for compounds that are involved in olfactory sensation. Modulators of
CC   their activity are useful for pharmacological and genetic modulation of
CC   olfactory signalling pathways. Therefore, they can be used in the food,
CC   pharmaceutical and cosmetic industries to customise odours and
CC   fragrances. The present sequence encodes a human olfactory receptor of
CC   the invention
XX
SQ   Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.58e-09      Length:          933
Score:                  20.00         Matches:         20
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            6.43%         Indels:          0
DB:                     5             Gaps:            0

US-10-023-597-24 (1-311) x AAS42233 (1-933)
```

```
Qy          79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291

RESULT 12
ABZ43067
ID   ABZ43067 standard; DNA; 933 BP.
XX
AC   ABZ43067;
XX
DT   06-MAR-2003  (first entry)
XX
DE   Human GPCR polynucleotide SEQ ID NO 395.
XX
KW   Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
KW   drug development; gustatory; taste; fragrance; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200216548-A2.
XX
PD   28-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-IB001446.
XX
PR   04-AUG-2000; 2000JP-00237818.
PR   13-FEB-2001; 2001JP-00034434.
XX
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI   Haga T,  Takeda S,  Mitaku S;
XX
DR   WPI; 2002-304118/34.
DR   P-PSDB; ABP95793.
XX
PT   Database global search for G protein-coupled receptors, proteins and
PT   encoded genes for studying in vivo signal transduction mechanism and
PT   identifying targets for drug development.
XX
PS   Claim 9; SEQ ID NO 395; 97pp + Sequence Listing; Japanese.
XX
CC   The invention relates to a method for screening G protein-coupled
CC   receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
CC   ABP95942) by extracting open-reading frames containing 6-8 transmembrane
CC   domains with 250-1000 amino acid residues to give a gene homologous with
CC   a known GPCR gene. The receptor proteins and encoded genes are useful for
CC   studying in vivo signal transduction mechanism and identifying targets
CC   for drug development e.g. based on olfactory and gustatory receptors in
CC   form of agonists and antagonists by screening intrinsic and extrinsic
CC   ligands as bitter taste inhibitors, taste enhancers and fragrance
CC   improvers. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.58e-09        Length:         933
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     6               Gaps:           0

US-10-023-597-24 (1-311) x ABZ43067 (1-933)

Qy          79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291

RESULT 13
ABK68455
ID   ABK68455 standard; DNA; 933 BP.
XX
AC   ABK68455;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Human DNA for olfactory and pheromone G protein-coupled receptor #55.
XX
KW   Human; ds; gene; olfactory and pheromone G protein coupled receptor;
KW   GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;
KW   anorectic; taste; fragrance; food additive; cosmetic; cell migration;
KW   sterility; psychotic disorder; neurological disorder; anxiety;
KW   schizophrenia; manic depression; depression; axonal growth;
KW   menstrual cycle; appetite sexual motivation; sexual attraction;
KW   aggression.
XX
OS   Homo sapiens.
XX
PN   WO200224726-A2.
XX
PD   28-MAR-2002.
XX
PF   21-SEP-2001; 2001WO-BE000162.
XX
PR   22-SEP-2000; 2000EP-00870211.
XX
PA   (CHEM-) CHEMCOM SA.
XX
PI   Veithen A;
XX
DR   WPI; 2002-330013/36.
DR   P-PSDB; AAU95568.
XX
PT   Novel pheromone G-protein coupled receptor and receptor-derived agonists,
PT   antagonists or inhibitors useful in food or cosmetic products or in the
PT   treatment or prevention of neurological disorders such as anxiety and
PT   schizophrenia.
XX
PS   Disclosure; Page 178-179; 833pp; English.
XX
CC   The invention relates to olfactory and Pheromone G-protein coupled
CC   receptor (GPCR) or a protein 95% identical to the GPCR, a specific active
CC   portion and its encoding polynucleotide. Also included are an agonist,
CC   antagonist or inhibitor of the GPCR or the polynucleotide, a vector
CC   comprising the polynucleotide, a cell transformed by the vector, a non-
CC   human mammal comprising a partial or total deletion of the polynucleotide
CC   encoding the receptor and screening (detection and possibly, recovering)
CC   of compounds which are known or not known to be agonist, antagonists or
CC   inhibitors of natural compounds to the GPCR. The receptor-derived
CC   agonists, antagonists, inhibitors or compounds are used as an
CC   improvement, elimination or substitution of an existing taste and/or a
CC   fragrance of (or in) the food and/or cosmetic products. They can also be
CC   used in the preparation of medicament in the treatment and/or prevention
CC   of a mammalian disorder, such as cell migration, sterility, psychotic and
CC   neurological disorders, including anxiety, schizophrenia, manic
CC   depression, depression, for promoting axonal growth, nerve cell
CC   connection and nerve regeneration for modulating male and female
CC   endocrine functions, hormone production and the menstrual cycle, for the
CC   prevention or the treatment by stimulation of several mammalian
CC   behaviours, such as stimulation or suppression of appetite, sexual
CC   motivation, sexual attraction, aggression and for promoting or
CC   suppressing chemical communication between organisms. The present
CC   sequence is a human DNA encoding an olfactory and pheromone GPCR
XX
SQ   Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.58e-09        Length:         933
Score:                  20.00           Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     6               Gaps:           0
```

US-10-023-597-24 (1-311) x ABK68455 (1-933)

```
Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291
```

RESULT 14
ABK37521
ID ABK37521 standard; cDNA; 933 BP.
XX
AC ABK37521;
XX
DT 08-MAY-2002 (first entry)
XX
DE DNA encoding G-coupled olfactory receptor #23.
XX
KW Human; olfactory G-coupled receptor; sensory perception of odourant;
KW odour composition; taste composition; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200198526-A2.
XX
PD 27-DEC-2001.
XX
PF 22-JUN-2001; 2001WO-US020122.
XX
PR 22-JUN-2000; 2000US-0213812P.
PR 13-MAR-2001; 2001US-00804291.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S, Stryer L;
XX
DR WPI; 2002-083330/11.
DR P-PSDB; AAU85162.
XX
PT Representing sensory perception of one or more odorants for the
PT identification and design of tastes and odors comprises providing a
PT representative group of n olfactory receptors.
XX
PS Example; Page 68; 182pp; English.
XX
CC The invention relates to a method of representing sensory perception of
CC one or more odourants. The method comprises: (a) providing a
CC representative class of n olfactory receptors or ligand binding domains
CC (LBD) of these receptors; (b) measuring values X1 to Xn representative of
CC at least one activity of one or more odourants selected from: (i) binding
CC one or more odourants to the LBD of at least one of the n olfactory
CC receptors; (ii) activating at least one of the n olfactory receptors with
CC the one or more odourants; and (iii) blocking at least one of the n
CC olfactory receptors with the one or more odourants; and (c) generating a
CC representation of sensory perception from the values X1 to Xn. The
CC representation of the sensory perception of odourants is useful for the
CC design and formulation of odour and taste compositions. ABK37499-ABK37754
CC and ABK37918-ABK37921 represent human olfactory G-coupled receptor coding
CC sequences and related PCR primers of the invention
XX
SQ Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.58e-09	Length:	933
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	6	Gaps:	0

US-10-023-597-24 (1-311) x ABK37521 (1-933)

```
Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 291
```

RESULT 15
ABT05655
ID ABT05655 standard; DNA; 953 BP.
XX
AC ABT05655;
XX
DT 16-OCT-2002 (first entry)
XX
DE GPCR 4 protein encoding DNA SEQ ID No 13.
XX
KW Antidiabetic; cytostatic; anorectic; nootropic; neuroprotective; GPCRX;
KW antiparkinsonian; cardiant; antiarteriosclerotic; immunosuppressive;
KW hypotensive; haemostatic; antifertility; antiasthmatic; antiinflammatory;
KW anti-HIV; G-protein coupled receptor X; cardiomyopathy; atherosclerosis;
KW diabetes; metabolic disorder; diabetes; obesity; infectious disease;
KW anorexia; neurodegenerative disorder; Alzheimer's disease; dyslipidaemia;
KW Parkinson's disorder; immune disorder; haematopoietic disorder; obesity;
KW metabolic syndrome X; wasting disorder; cancer; hypertension; neoplasm;
KW congenital heart defect; aortic stenosis; subaortic stenosis; lymphoma;
KW transplantation; adrenoleukodystrophy; congenital adrenal hyperplasia;
KW prostate cancer; adenocarcinoma; uterus cancer; fertility; haemophilia;
KW hypercoagulation; idiopathic thrombocytopenic purpura; immunodeficiency;
KW graft versus host disease; AIDS; bronchial asthma; Crohn's disease;
KW multiple sclerosis; chromosomal mapping; tissue typing; forensic biology;
KW gene therapy; transgenic animal; gene; ds.
XX
OS Unidentified.
XX
PN WO200246229-A2.
XX
PD 13-JUN-2002.
XX
PF 05-DEC-2001; 2001WO-US046530.
XX
PR 05-DEC-2000; 2000US-0251459P.
PR 29-DEC-2000; 2000US-0259007P.
PR 04-DEC-2001; 2001US-00005041.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Casman SJ, Padigaru M, Burgess CE, Shimkets RA, Spytek KA;
PI Gilbert JA, Mayotte JE, Baumgartner JC, Mishra V, Vernet CAM;
PI Dickinson KS, Ballinger RA, Wolenc AR;
XX
DR WPI; 2002-537559/57.
DR P-PSDB; ABJ04724.
XX
PT Novel isolated G-protein coupled receptor polypeptide, designated NOVX,
PT useful for treating or preventing in human receptor-associated disorders
PT e.g. cardiomyopathy, atherosclerosis or diabetes.
XX
PS Claim 9; Page 33; 264pp; English.
XX
CC The invention relates to a novel isolated G-protein coupled receptor X
CC (GPCRX) polypeptide. The isolated protein, its encoding polynucleotide
CC and the antibody of the isolated protein is useful for treating or
CC preventing a GPCRX-associated disorder in a subject, preferably human,
CC where the disorder is cardiomyopathy, atherosclerosis or diabetes. The
CC isolated protein, its encoding polynucleotide and the antibody of the
CC isolated protein is useful for treating or preventing metabolic
CC disorders, diabetes, obesity, infectious disease, anorexia,
CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC immune disorders, haematopoietic disorders, and various dyslipidaemias,
CC metabolic disturbances associated with obesity, the metabolic syndrome X,
CC wasting disorders associated with chronic diseases, and cancer. The
CC disorders also include cardiomyopathy, atherosclerosis, hypertension,
CC congenital heart defects, aortic stenosis, subaortic stenosis,
CC transplantation, adrenoleukodystrophy, congenital adrenal hyperplasia,
CC prostate cancer, neoplasm, adenocarcinoma, lymphoma, uterus cancer,
CC fertility, haemophilia, hypercoagulation, idiopathic thrombocytopenic
CC purpura, immunodeficiencies, graft versus host disease, AIDS, bronchial

CC asthma, Crohn's disease, and multiple sclerosis. The isolated protein,
CC its encoding polynucleotide and the antibody of the isolated protein is
CC useful in screening assays, detection assays (e.g., chromosomal mapping,
CC tissue typing, forensic biology). The isolated polynucleotide is useful
CC in gene therapy, to express the isolated protein, to detect GPCRX mRNA or
CC a genetic lesion in a GPCRX gene, and to modulate GPCRX activity. The
CC cell of the invention is useful for producing non-human transgenic
CC animals. This polynucleotide sequence represents DNA encoding a GPCRX
CC protein of the invention
XX
SQ Sequence 953 BP; 202 A; 256 C; 193 G; 302 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.64e-09	Length:	953
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	6	Gaps:	0

US-10-023-597-24 (1-311) x ABT05655 (1-953)

```
Qy         79 ProLysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 CCCAAAATGCTGATGAGTTTTGTCTCAAGGAAGAACATCATTTCCTTCACAGGGTGTATG 306
```

Search completed: September 30, 2004, 14:02:34
Job time : 451 secs

This Page Blank (uspto)